DOCUMENT RESUME

ED 217 169 CE 032 584

AUTHOR Sticht, Thomas G.; And Others
TITLE Literacy, Oracy, and Vocational Aptitude as Predictors of Attrition and Promotion in the Armed Services.
INSTITUTION Human Resources Research Organization, Alexandria, Va.
SPONS AGENCY Office of the Assistant Secretary of Defense for Manpower and Reserve Affairs (DOD), Washington, D.C.
REPORT NO HumRRO-PP-2-82
PUB DATE Mar 82
CONTRACT MDA903-80-C-0588
NOTE 82p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Attrition (Research Studies); *Listening Comprehension Tests; Listening Skills; *Literacy; Military Training; *Oral Language; Postsecondary Education; Predictive Measurement; *Predictive Validity; Promotion (Occupational); Reading; Reading Tests; Speech Skills; Test Reliability; Test Validity; *Vocational Aptitude
IDENTIFIERS Armed Services Vocational Aptitude Battery; *Literacy Assessment Battery

ABSTRACT

A project was conducted to (1) produce normative data for the Literacy Assessment Battery (LAB) and (2) evaluate the LAB as a potential supplement to the Armed Services Vocational Aptitude Battery (ASVAB) for use as a selection and classification instrument for the military services. The distribution of auding and reading skills in the population that applies for military service was first determined by administering the LAB test to more than 4,500 applicants for service; these scores were normed and related to other literacy tests and composite scores of the ASVAB. It was found that auding and reading are highly correlated, indicating that people who are unskilled at reading are the least skilled in comprehending oral language also, However, for the lowest-scoring populations, the lower the reading skill, the more likely the auding skills were better. The predictive validity study showed the value of the LAB for predicting qualification status, predicting attrition, and predicting promotion (although education level emerged as the best single predictor of promotion). It was concluded that understanding the nature of literacy and the relationships of literacy to aptitude assessment can lead to improvements in selection and classification. The LAB should continue to be used in conjunction with the ASVAB to predict success in training schools where the demands for literacy and oracy skills are higher than on the job. (KC)

Professional Paper 2-82

HumRRO-PP-2-82

ED217169

March 1982

HumRRO

Literacy, Oracy, and Vocational Aptitude as Predictors of Attrition and Promotion in the Armed Services

Thomas G. Sticht
Lydia R. Hooke
John S. Caylor

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street • Alexandria, Virginia 22314

PREFACE

The present report is one of a series resulting from research under Contract No. MDA903-80-C-0588 Evaluation of the Predictive Validity of the Literacy Assessment Battery (LAB). This research is sponsored by the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics), Directorate for Accession Policy. Dr. W.S. Sellman is the technical monitor for this work. Data for norming the LAB were provided by Dr. Lonnie Valentine and Mr. John Mathews, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas. Data analysis support was provided by the staff of the Defense Manpower Data Center (DMDC) in Alexandria, Virginia and Monterey, California. The support of these individuals and organizations in the conduct of this research is greatly appreciated.

EXECUTIVE SUMMARY

BACKGROUND

As a part of a continuing program of research to improve selection and classification testing in the military, the OASD(M&RA) sponsored research during Project 100,000 that indicated that the assessment of auding (listening) skill contributed significantly to the ability to identify the more capable job performers among personnel in Mental Category IV. Subsequent work, sponsored by the Air Force, produced a theoretical model of the relationship of auding to reading skills that suggested reasons why auding assessment might contribute to the prediction of job performance among lower-aptitude personnel.

There were two reasons: first, many lower-aptitude persons are unskilled readers, and the ASVAB tests, which demand reading skills for successful performance, cannot discriminate, among the poor readers, between those with poor oral language comprehension skills and those with relatively well-developed oral language skills but who have problems with written language. An auding test makes the detection of such persons possible. Second, job skills and knowledges can often be learned through watching and listening. Reading skills may play a minor role in the accomplishment of many job tasks, but auding comprehension is required for performing even the most routine tasks (e.g., verbal orders must be followed). Auding assessment helps identify persons with reasonably well developed oral language comprehension skills who can learn to perform job tasks through observation and listening.

The theoretical work on auding and reading suggested that, in the typical case, reading skills are based upon preexisting ability to comprehend oral language. Preschool children and illiterate adults generally have some degree of oral language skill prior to learning to read. Then, in learning to read, they close the gap between oral and written language comprehension abilities. Based on these ideas, the Air Force Human Resources Laboratory sponsored additional research that produced a Literacy Assessment Battery (LAB) designed to assess the "gap" between a person's auding and reading skills. However, the LAB test was developed only in an experimental form, and was not normed on a sample of young adults like those who apply for military service, nor was the effectiveness of the LAB as a selection and classification tool evaluated.

PROBLEM

The present research was conducted to (1) produce normative data for the LAB, and (2) evaluate the LAB as a potential supplement to the Armed Services Vocational Aptitude Battery for use as a selection and classification instrument.

APPROACH

To determine the distribution of auding and reading skills in the population that applies for military service, Mathews, Valentine, and Sellman (1978) administered the LAB test to over 4500 applicants for service as a part of research to study reading skills of applicants for military service.

The present research analyzed the LAB data obtained by Mathews et al., to develop normative data for the LAB, and to relate scores on the LAB to other literacy tests and composite scores of the ASVAB. Additionally, the LAB and other literacy and aptitude data were merged with data on qualification status, attrition, and paygrade attained contained in the computer data files of the Defense Manpower Data Center (DMDC). The latter data served as the criteria for evaluating the predictive validity of the LAB and the AFQT composite from the ASVAB.

MAJOR FINDINGS

(1) Normative data regarding oracy and literacy skills. Consistent with the theoretical model it was found that:

- Auding and reading are highly correlated (.73 for LAB Auding and Reading Paragraphs) indicating that people who are unskilled at reading are the least skilled in comprehending oral language, also.
- The lower the reading skill, the more likely one is to find people who comprehend better by auding than by reading; however, overall, across the full range of reading skills in the norming population, except the very lowest levels, people tended to perform better by reading than by auding. This suggests that for the majority of poor readers who apply for military service, literacy problems are accompanied by oracy problems (low oral language vocabulary and low ability to process orally presented material).

(2) Predicting qualification status, attrition, and promotion. The predictive validity study obtained results showing the value of the LAB for:

- Predicting qualification status. The LAB total score correlated almost as well with Qualification status (qualified vs. not-qualified for military service) as did the AFQT, which was actually used in determining the qualification status, reflecting to a large extent, the fact that both the LAB and AFQT assess language and reading knowledge and skills (AFQT and LAB Total scores correlate .71).
- Predicting attrition. The prediction of attrition from military service was performed separately for attrition in the first six months, which reflects the learning demands of initial entry and job training and orientation, and months 7 through 30, which reflects a period of job performance in which cognitive skills are stressed least. Additionally, months to attrition was used as a criterion for that subpopulation which in fact attrited. Results supported the hypothesis that adding oracy skills to the literacy skills assessed by the AFQT would improve the accuracy of selection and classification procedures. Though in no case were correlations very large, none-the-less in three out of the four predictive validity evaluations, the LAB Auding Paragraphs subtest emerged as the best LAB subtest to supplement the AFQT, and increased the validity of the AFQT two or threefold in predicting attrition from the military in the first 30 months.

- Predicting promotion (paygrade achieved). Education level emerged as the best single predictor of paygrade achieved, no doubt reflecting the practice across the Services of using education level as one criterion for promotion to higher rank. The LAB Auding Paragraphs test once again emerged as the best LAB subtest to supplement the AFQT in predicting paygrade.

CONCLUSIONS

Understanding the nature of literacy and the relationships of literacy to aptitude assessment can lead to improvements in selection and classification. Continued exploration of the Literacy Assessment Battery (LAB) as a selection and classification supplement to the ASVAB should be undertaken to predict success in training schools where the demands for learning by oracy and literacy skills are higher than on-the-job. For training settings, criterion measures, for example, end-of-course grades, are more likely to reflect the types of cognitive skills assessed by the LAB and ASVAB subtests and composites than do the behavioral factors that make up the majority of the reasons given for attrition from the military. With a greater correspondence among the skills and knowledges represented in the predictor and criterion variables, predictive validity should increase.

TABLE OF CONTENTS

Page

TABLE		Page
13	Intercorrelation Among LAB, Reading Test, and ASVAB Subtests	37
14	LAB and ASVAB Intercorrelation Matrix Interpreted in Accordance With the Developmental Model of Literacy	39
15	Criterion Variables Used in Predictive Validity Study	43
16	Reasons for Attrition (Interservice Separation Codes) During and Subsequent to First Six Months of Service	45
17	Description of Populations Involved in Predictive Validity Study	48
18	Coefficients of Predictive Validity (r) for LAB and AFQT Scores	50

Literacy, Oracy, and Vocational Aptitude as Predictors of Attrition and Promotion in the Armed Services

Chapter 1

INTRODUCTION

Applicants for military service must meet the standards for enlistment established by Congress, the Department of Defense and each of the armed services. In addition to educational selection standards, applicants must also meet the criteria established for performance on the Armed Services Vocational Aptitude Battery (ASVAB).

However, even though the ASVAB is at the state-of-the-art as a personnel selection and job classification instrument, there are indications that some applicants whose performance on the ASVAB disqualifies them from service could, if given the opportunity, successfully complete a job training program and be productively utilized in a duty position. This was evidenced in the late 1960s when, under Project 100,000, several hundred thousand individuals were enlisted whose aptitude scores were well below the standards then in existence. While, as a group, these "new standards" personnel did not learn as well (Fox, Taylor, and Caylor, 1969), and did not perform as well on job knowledge and job sample performance tests as did higher aptitude groups (Vineberg, Sticht, Taylor, and Caylor, 1971), literally tens of thousands learned and performed as well as service members whose aptitude scores met the higher standards of the pre- and post-Project 100,000 periods.

APTITUDE AND LITERACY IN SELECTION AND CLASSIFICATION

The Project 100,000 research, and numerous other studies conducted by the various research and development offices of the armed services (see the annual proceedings of the Military Testing Association), have aimed to improve the selection and classification accuracy of the ASVAB to more effectively draw upon and utilize what is becoming a shrinking work force. Military manpower analysts report that population figures indicate a reduction in the population of people aged 17 to 24 from 13% in 1975 to 8% in 1985 (Canter, 1978). It is, therefore, in the interest of the armed services to find ways to assess more carefully the capabilities of this shrinking youth population so that selection and classification decisions are not wasteful of scarce human resources.

Compounding the problems arising from a reduction in the youth population from which the military typically obtains new personnel are reports that the literacy (verbal and mathematics) skills of this shrunken population are lower today than during the mid-sixties and earlier (Waters, 1981). While the operational consequences of this reported decline are uncertain, the new forms of the ASVAB (Forms 8, 9, and 10) implemented in October 1980 were designed to (1) more accurately estimate ability, particularly in the lower test score range, and (2) provide a better measure of literacy (Report to the House and Senate Committees on Armed Services, OASD(MRA&L), December 31, 1980).

The latter report goes on to indicate further the importance of literacy, in the form of reading, to the selection and classification process. First, with regard to the ASVAB itself, the report notes that ". . . portions of the new ASVAB were designed to provide a better measure of reading skills than tests previously used. Because ASVAB 8, 9, and

10 have a high degree of correlation with commercial reading tests, they improve DoD's ability to screen out applicants with marginal literacy skills." (p. 7). Table 13 of this report shows that correlations of ASVAB Forms 6 and 7 with general reading are in the .66 to .79 range. Mathews, Valentine, and Sellman (1978) reported a multiple correlation between three ASVAB subtests (Word Knowledge, General Science, Numerical Operations) and reading of .86 (p. 10).

The importance of reading to selection and classification is further evidenced through direct examination of sample ASVAB items that are available for public examination. For ASVAB Form 5, there are 12 group-administered tests of which 10 require reading of words, sentences, numerals and arithmetic operator symbols (+, -, x, ÷). Thus, reading is a component of practically all of the subtests comprising the ASVAB Form 5, and the new ASVAB increases this use of reading by dropping the Space Perception subtest and including the Paragraph Comprehension subtest.

The OASD(MRA&L) report to Congress also points out that, whereas in the past, the ASVAB has been primarily validated as a predictor of trainability, and hence it has been validated against criteria of success in training, there is a new effort to validate the ASVAB as a predictor of job performance. For training programs, and increasingly for job performance, the criteria of success have included performance on paper-and-pencil tests that require literacy skills. For instance, the Army has initiated the use of Skill Qualification Tests (SQTs) which assess job proficiency using both written and hands-on performance components. Research that influenced the policy decision to develop SQTs indicated that, averaged over four Army jobs, general reading ability correlated about .50 with the paper-and-pencil test component and .33 with the hands-on component (Sticht, 1972, p. 291). More recent data correlating operational SQT performance with general reading confirm these earlier findings (Miller, Nystrom, and Hicks, 1980).

In summary, it is clear that literacy is a concept that is related to the business of selection and classification in DoD. First, there is concern that the literacy skills of the youth population from which new recruits are drawn are declining. Second, the ASVAB requires reading as a necessary but not sufficient information processing skill for performance beyond random guessing. Third, the ASVAB is more and more being validated as a selection and classification instrument against job proficiency criteria that contain literacy skills as components for performance. Paper-and-pencil tests which were once the criteria for validating ASVAB tests in the training base against end-of-course grades, are now being developed to assess job proficiency at the permanent duty station and, increasingly, are being used as a factor in determining promotion to higher paygrades (Army Regulation 600-200, 1 January 1981). Literacy thus becomes a concept that is common across predictor and criterion measures.

Given the centrality of literacy in the development and validation of the ASVAB as the DoD's primary selection and classification tool, it is important to have as clear an understanding as is possible regarding (1) the nature of literacy and the role of literacy in aptitude assessment, and (2) the implications of such knowledge for increasing the size of the utilizable work force for military accessions and the reduction of attrition through the more accurate matching of the capabilities of recruits to the requirements of military training programs and jobs.

Chapter 2 of this report presents an analysis of literacy and discusses the role of literacy in aptitude assessment.

Chapter 3 summarizes research to develop a Literacy Assessment Battery that was based on the analysis of literacy outlined in Chapter 2.

In Chapter 4, a study is reported that explores the use of the Literacy Assessment Battery as a selection and classification supplement to the ASVAB. While this research was conducted using ASVAB Forms 6 and 7, and is therefore not directly applicable to the new ASVAB 8, 9, and 10, the research is useful as an exploratory project to determine if further developments along this line of inquiry are warranted.

Finally, Chapter 5 presents a discussion of the findings and concepts presented in Chapters 2, 3, and 4 and draws implications of the present research for understanding the concept of "aptitude" and the assessment of aptitude for more effective selection and classification.

Chapter 2

THE NATURE OF LITERACY

Intelligence, IQ, aptitude, and literacy are concepts that are intertwined in ordinary parlance. For instance, a September 1980 article in *Science* magazine shows how these concepts merge in common usage:

> "Although the Army,[1] in common with everyone else who attacks aptitude and IQ tests, insists that verbal skills and ability in handling written tests are not necessarily indicators of how a person will actually function, this is a debatable proposition. Literacy in the Army is notoriously low. For example, a check at Fort Benning in 1976 revealed that 53 percent of the enlistees had a 5th grade or lower reading ability. There are many experts who regard literacy not just as one of many desirable skills, but as a fundamental indicator of mental ability." (Holden, 1980, p. 1099)

But it is not only in semi-popular writing that the conceptual commingling of literacy and aptitude, intelligence, etc., is evidenced. In a technical report describing correlations of reading test performance with Skill Qualification Test (SQT) scores for several military occupational specialties (MOS), the following discussion is found of correlations among the written and hands-on components of the SQT:

> "The most important correlations are between Written Percent Correct (WPCT) and reading test scores (.41, .40, and .46, for word knowledge, reading comprehension and reading level, respectively). This indicates a substantial relationship between reading ability and WPCT.
>
> However, part of these correlations may be explained by a "g" factor ("general ability" or "general intelligence") that relates to both written and hands-on tests, rather than a special reading factor.
>
> . . . (hands-on percent correct) correlated about as much with the reading test as with the WPCT, indicating some sort of "g" factor underlying all three tests." (Miller, Nystrom, and Hicks, 1980)

Here, some sort of "general intelligence" factor is postulated to be "underlying" the reading tests and written and hands-on components of the SQT, again illustrating the interrelatedness among the concepts of literacy (reading), aptitude, and intelligence.

It seems clear, from the foregoing, that there is a need to disentangle the concepts of intelligence, aptitude, verbal abilities, and literacy so that more appropriate assessment of work force capabilities can be achieved. This is a more pressing concern when it is recognized that literacy, but not aptitude or intelligence, is thought to be teachable and learnable. Clearly, there would be implications for defense accession policies, education, and training if it should turn out otherwise.

[1] This quote does not necessarily represent the official position of the Department of the Army.

A DEVELOPMENTAL MODEL OF LITERACY

Because of the centrality of the concept of literacy for selection, classification, training, and job performance in the military, the OASD(M&RA), in 1968, initiated research by the Human Resources Research Organization on the nature and uses of literacy in military environments (Sticht, 1975). Early on in the conduct of that research, a rudimentary conception of the hierarchical, developmental relationships among certain basic adaptive processes, oral language, and reading was articulated (Sticht, 1972). This work was continued under the sponsorship of the Air Force Human Resources Laboratory and resulted in the publication of a more fully developed model of the relationships among basic information processing skills, oral and written language and the contents of memory that are used to express or comprehend knowledge presented in oral or written language (Sticht, Beck, Hauke, Kleiman, and James, 1974).

The goal of the effort to construct a developmental model of literacy is to explain, at a macro-level, how children who are born, as illiterates, into our literate culture, acquire the knowledge and skills required to perform the vast numbers of literacy tasks that our society has come to expect citizens to perform. This includes, of course, the kinds of literacy tasks posed by the ASVAB or any test of intelligence, aptitude, or achievement that requires the use of graphic (printed; written) language and non-language (e.g., →) symbols.

Figure 1 presents, in schematic form, an overview of the major concepts included in the developmental model of literacy. Before addressing the details of the model, several orienting comments regarding Figure 1 are in order. First, the figure is meant to portray a developmental sequence when examined from left to right. The sequence begins with the newborn infant, and goes through stage 4 in which literacy skills are functional. The broad arrowhead on the far right is meant to imply continued development over the lifespan. The development of literacy, language, and knowledge is a lifetime activity.

Examining Figure 1 from top to bottom, the top series of boxes are meant to represent the environment in which the person exists. This is the environment "outside the head." This external environment makes available information displays in the form of structured energy (mechanical for hearing; electromagnetic for seeing) that the person can explore and transform into internal representations of the external information. These internal representations are developed by the processes in the second series of boxes labeled, on the far left, Information Processes in Working Memory. These processes go on "inside the head," and merge information picked-up from the external information displays with information picked-up from the third series of boxes, labeled on the far left as Long Term Memory. Thus, the processes in the working memory are used to pick-up and merge information from outside the brain with information in long term memory inside the brain to construct an internal representation of the world as currently experienced, including the meaning of symbolic information when this is a major domain of information being extracted from the external world at a given time.

At the top of Figure 1, there are references to four stages. Stage one refers to the newborn infant who is considered to be innately endowed with the Basic Adaptive Processes involved in sensory/perceptual processes such as hearing and seeing, etc., motor movement, and cognition, including the processes needed to acquire information, mentally manipulate it, store it in memory, form knowledge structures out of it (e.g., images, facts, concepts, principles, rules, and, after considerable experience in life, large structures of subject matter areas such as mathematics, MOS knowledge, etc.) and retrieve and represent the information in various ways. In stage 1, these processes are assumed to work more or less automatically without conscious control.

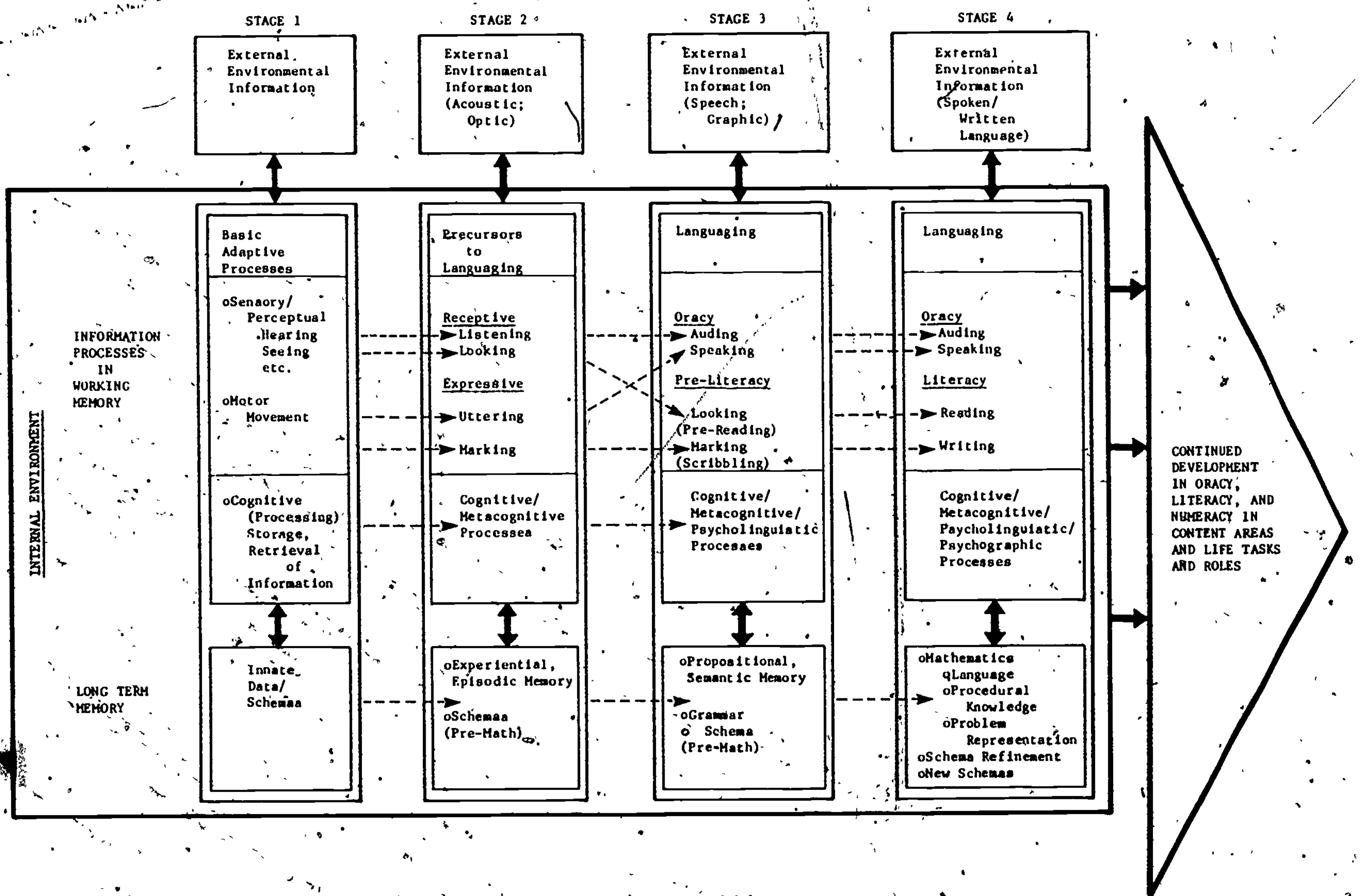


Figure 1. **A Developmental Model of Literacy**

Stage 2 represents the emergence of conscious control over information pick-up and manipulation. This active process of attending to information distinguishes listening from hearing, and looking from seeing, as information pick-up processes. Listening and looking build internal representations that may be called images (though at times this general term is divided into echoic and iconic images for listening and looking, respectively). Images may also be constructed from data stored in long term memory. These internal imaging processes are frequently assessed in aptitude tests as "spatial perception" or "mechanical comprehension" in which it is necessary to mentally visualize and rotate cog-and-gear assemblies to determine what effect this movement might have on some other gear.

Stage 2 also introduces the concept of active or working memory, which is defined by the occurrence of consciously controlled information processing activities. Working memory is a limited memory that can easily be overloaded (e.g., attending to 2 or 3 things at once is difficult—if not impossible). Many of the information processing activities the person acquires will be techniques to overcome active memory limits (e.g., repeating information to oneself keeps the information in active working memory until it can be applied).

Stage 3 represents the development of language processes out of earlier processes and knowledge stored in long-term memory. In developing oral language, the listening process is used in attending to spoken language to learn the words and grammar of language. Thus, listening plus languaging, occurs simultaneously. This joint occurrence is given the special name of auding. On the production side, the joint occurrence of uttering (making sounds through the mouth) with the production of word forms from the language pool, and stringing the word forms together to make sentences using the rules of grammar, produces the special process called speaking. Auding and speaking comprise the oral language information reception and production skills. Speaking is used to represent information that the person has in his or her mind "outside the head" and in the acoustic medium, while auding is used to pick-up and decode speech information displays into knowledge in the mind of the listener.

In stage 4, the information processing skills of looking and marking are used to learn a representational system which, in many respects, represents the spoken language in a different medium—light—and in a more or less permanent graphic display: the written language. Looking at written language and transforming the written language into meaning is called reading. Writing is the special use of marking skills to produce graphic language (and other symbols and symbol systems).

In the typical case, people develop a fair amount of competence in oral language before they are exposed to formal instruction in reading in the elementary grades. The written language skills build upon the earlier developed oral language skills and add new vocabulary and concepts, as well as special knowledge about how to represent information in the graphic medium, to the person's knowledge base. In turn, learning new vocabulary and conventions of language through reading and writing enlarges the person's oral language abilities. The large arrow at the far right in Figure 1 is meant to represent the notion that the development of oral and written language ability may continue indefinitely. As mentioned earlier, learning to read can take a lifetime.

A major component of Figure 1 is the person's long-term memory. The long-term memory contains all the knowledge developed by the person in interaction with the environment, including the processes the brain invents to overcome limitations in working memory and other aspects of its functionings (such as retrieval processes for remembering information). Much of the knowledge acquired by the person will never be understood in consciousness (for example, the rules of grammar). Rather, it will be unconsciously used to accomplish tasks such as developing language competency and comprehending the events of the world. In addition to the general world knowledge and processes, that are in the mind, though not accessible to conscious understanding without considerable analysis, the memory also contains the language knowledge (words and grammar) that can be used to represent information that arises from experience in the world (e.g., bodies of knowledge about parts of the body, houses, neighborhoods—sometimes called "schema" in cognitive science terms) and from didactic instruction, as in training programs.

A major body of knowledge frequently represented in aptitude and literacy tests is mathematics. This body of knowledge contains information acquired early in life, e.g., concepts such as "more than," "less than," "bigger than," "more of," and facts, procedures, concepts, principles, and rules explicitly taught in school and organized into large schemas addressed by labels such as algebra, geometry, trigonometry, statistics, etc. Much of the mathematics knowledge is language knowledge and notational system knowledge. However, because mathematics does not draw upon a new medium (i.e., it uses the acoustic and optic media used in oral and written language and in imagery), it does not require special information pick-up and representation processes. It utilizes both oral and written language representation systems, and expands these through the development of special vocabulary, rules for manipulating sentences (e.g., logic), special symbol/notational systems, and special graphic displays (graphs; charts) to represent mathematics concepts. One looks at and reads such displays and produces them by writing and marking (drawing), and one talks about them and auds lectures about mathematics; but no new processes for encoding or decoding information into or from acoustic or optic displays in the environment are required beyond those used in oracy and literacy tasks.

The foregoing and Figure 1 briefly summarize the structure of the developmental model of literacy and emphasizes: (1) three major components of knowledge in memory, processing skills, and the environment which provides a source of information, and (2) the development of later skills and knowledges as extensions to or transformations of earlier acquired skills and knowledges. The latter processes include the development of oral language skills from the earlier prelinguistic content and the processes of listening and uttering, and the written language skills of reading and writing as specializations of looking and marking coupled with earlier acquired language knowledge developed as oracy skills. This permits reading and writing to serve both as a second signaling system for representing oral language and as language/symbol systems in their own right that have created special information structures and displays unique to the visual medium: lists, charts, forms; organizations such as paragraphs, chapters, etc. These ideas, only briefly touched on here, are elaborated in various reports (Sticht & Beck, 1976, Sticht, 1977, 1978).

AN APPLICATION OF ONE ASPECT OF THE DEVELOPMENTAL MODEL TO LITERACY AND APTITUDE ASSESSMENT

In the developmental model, before they learn to produce or to read written language, children develop oral language skills and skills in interpreting visual signs, such as facial gestures. Further, they develop skill in making graphic marks for drawing, scribbling, coloring, and so forth. Thus, before they are literate in written language, children begin to use visual information displays in a communicative manner, and they develop skill in signaling their thoughts to others through the manipulation of the acoustic medium, that is, through comprehending and producing speech.

A unique feature of the written language is that it has evolved over centuries as a graphic means of representing language signaled first by speech. With the invention of the alphabet, a technology has been provided by means of which thoughts that can be signaled by spoken words can be signaled by a graphic representation of the spoken word. Thus, the written language is said to provide a second signaling system for speech (Fries, 1963).

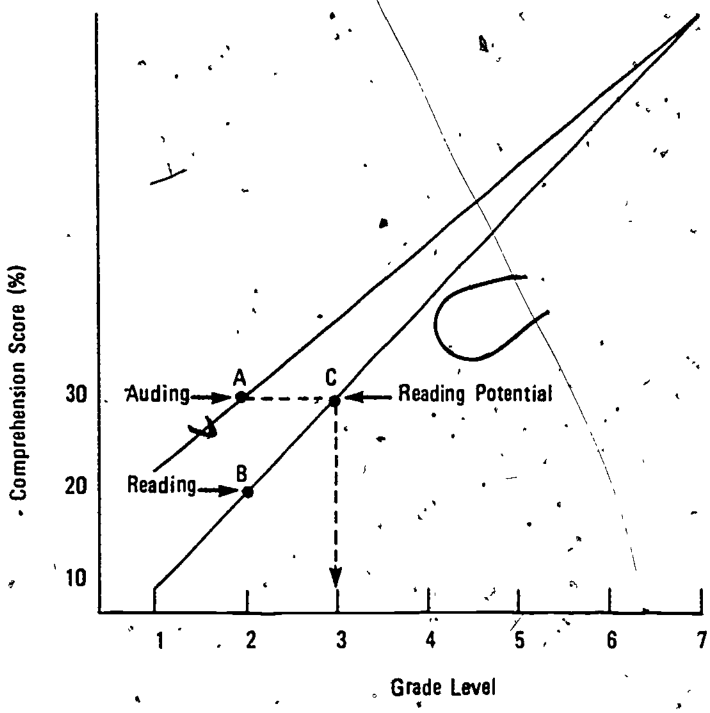
The use of written language as a second signaling system for speech is a continuation and extension of the child's symbolic use of visual, graphic information displays in the pre-language and oral language stages of development. Only now the graphic displays must be understood as standing for words that are already known in the spoken language. In learning to read, then, a major task is to learn to comprehend the printed form of language with the same accuracy and efficiency as one can comprehend language.

Because people typically learn to comprehend language by auding before they can comprehend it by reading, it is possible to consider that, in learning to read, they close the "gap" between the earlier achieved auding skill and the later developed reading skill which permits them to now comprehend in printed language what they could previously comprehend only in the spoken language. This process is illustrated schematically in Figure 2, where it is seen that, at the beginning of schooling, children comprehend language better by auding than by reading. However, as they progress through the school grades and acquire more and more skill in reading, they close the gap that existed between auding and reading skills at the outset of schooling.

A second way of regarding the auding-reading gap concept is to consider a situation in which a child in the second grade is presented the task of comprehending a brief written story and does quite poorly. Because the written language is a second signaling system for speech, the hypothetical data of Figure 2 suggest that if we present the story in spoken form, the child will comprehend it better because the gap between auding and reading has not yet been closed and auding exceeds reading skill in the second grade.

Following a comparable line of thought, it is possible that young adults who are very unskilled readers may perform poorly on the ASVAB because it presents all tasks in written language. If these poor readers actually comprehend better by auding than by reading, then a more accurate estimate of what they know and can communicate, comprehend, and learn might be obtained by means of an oral language test than by use of the written language test. In other words, because the written language is a second signaling system for speech, it is possible to remove the literacy component from aptitude testing by use of the primary language skills of speaking and auding. This makes it possible to obtain estimates of the extent to which low aptitude test performance reflects poor reading skills. In effect, this is a method of componential analysis similar to those that have been applied to reading (Frederiksen, 1981), aptitude (Pellegrino and Glaser, 1980), and intelligence (Sternberg, 1971) testing, albeit at a more macroscopic level.

Chapter 3 describes the development and norming of a literacy assessment battery that was designed based on the concepts discussed above and illustrated in Figure 2. In Chapter 4, a study is described in which the experimental literacy test battery is evaluated as an aptitude measure in terms of its effectiveness in predicting success in the military.



A – Indicates the normative auding score for the 2nd grade, called auding at the 2nd grade level.

B – Shows the normative reading score for the 2nd grade, called the 2nd grade level.

C – Shows conversion of the normative auding score to a reading "potential" score by drawing a horizontal from A to intersect with the reading curve, and then dropping a perpendicular line to the abscissa.

The example shows a reading potential score of 3rd grade.

Thus, the case illustrated shows a person auding and reading at the 2nd grade level, with a reading potential score of 3rd grade level.

Figure 2. **Schemata Showing Relationships Among Auding and Reading Comprehension Scores as a Function of School Grade Level**

Chapter 3

DEVELOPMENT OF THE LITERACY ASSESSMENT BATTERY (LAB)

As a part of a continuing program of research to improve selection and classification testing in the military, the Office of the Assistant Secretary of Defense (Manpower and Reserve Affairs) sponsored research during Project 100,000 that indicated that the assessment of auding (listening) skill contributed significantly to the ability to differentiate among personnel in Mental Category IV (Sticht, Kern, Caylor and Fox, 1970). Subsequent work sponsored by the Air Force Human Resources Laboratory, produced the theoretical model of the relationship of auding to reading skills discussed in Chapter 2 (Sticht, Beck, Hauke, Kleiman and James, 1974). This model suggested reasons why auding assessment might contribute to the prediction of job performance among lower-aptitude personnel.

There were two reasons: first, many lower-aptitude persons are unskilled readers, and the ASVAB tests, which demand reading skills for successful performance, cannot discriminate among the poor readers, between those with poor oral language comprehension skills and those with relatively well-developed oral language skills but who have problems with written language. An auding test makes the detection of such persons possible. Second, job skills and knowledges can often be learned through watching and listening. Reading skills may play a minor role in the accomplishment of many job tasks, but auding comprehension is required for performing even the most routine tasks (e.g., verbal orders must be followed). Auding assessment helps identify persons with reasonably well developed oral language comprehension skills who can learn to perform job tasks through observation and listening.

The theoretical work on auding and reading demonstrated that, in the typical case, reading skills are based upon preexisting ability to comprehend oral language. Preschool children and illiterate adults generally have some degree of oral language skill prior to learning to read. Then, in learning to read, they close the gap between oral and written language comprehension abilities (see Figure 2). Based on these ideas, the Air Force Human Resources Laboratory sponsored additional research that produced a Literacy Assessment Battery (LAB) designed to assess the "gap" between a person's auding and reading skills (Sticht and Beck, 1976).

The major purpose of the LAB test is to assess the degree to which the "gap" between auding and reading has been closed. Additionally, the LAB contains subtests that aim to clarify performance by auding and by reading. The various tests comprising the LAB are summarized in Table 1.

The Literacy Assessment Battery is comprised of three tests:

1. A Paragraphs Test designed for comparing how well an adult reads connected text to how well he or she auds comparable materials under comparable conditions, e.g., whether or not there is a "gap" between comprehension by these two modalities. In the Paragraphs Test, two passages are read aloud to examinees (Auding Subtest), and they are asked to read an additional two passages to themselves (Reading Subtest). These passages consist of adult-oriented materials, written at approximately the 9th grade level of difficulty according to the FORCAST readability formula (Sticht, 1975),

Table 1

Components of the Literacy Assessment Battery (LAB)

Test	Material	Task	Presentation Mode	Time/Rate	Total Time to Administer	Maximum Score
Paragraph:						
Auding	Two adult-oriented passages, 150 and 190 words long: FORCAST RGL* of 9th grade.	To answer 12 constructed response questions per passage involving recall of facts.	Passages & questions read aloud by examiner. (Auded by examinee)	1 minute, 15 seconds and 1 minute, 30 seconds per passage. 2 minutes and 30 seconds per 12 questions.	10	24
Reading	Same as above	Same as above	Passages & questions read silently by examinee.	Same as above	10	24
Vocabulary:						
Auding/ Reading	14 words from each passage presented in 3-4 word context from passage.	To choose correct synonym for each word from among 4 alternatives. (multiple choice)	Questions and alternatives simultaneously read and auded.	7 minutes	10	28
Reading	Same as above	Same as above	Questions and alternatives read silently.	Same as above	10	28
Decoding:						
I	Adult-oriented passage about 330 words in length: FORCAST RGL of 9th grade.	To detect and circle mismatches between words read on page and words read aloud by examiner.	Simultaneously read and auded.	100 words per minute.	3	10
II	Same as above	Same as above	Same as above	150 words per minute.	3	10
III	Same as above	Same as above	Same as above	200 words per minute.	3	10
IV	Same as above	Same as above	Same as above	250 words per minute.	3	10
					LAB TOTAL	144

*i.e., written at ninth grade reading level as measured by the FORCAST formula for assessing readability.

which were empirically equated for difficulty in a calibration study. Within each modality, one passage gives a sequence of procedures for performing a task and the other is descriptive. Passages for the two modalities are of matched length (150 and 190 words) and reading time allowed in the Reading Subtest is the same as the time it takes to read the Auding Subtest passages aloud. Times were equated so that auding and reading skills could be compared in terms of efficiency as well as accuracy.

After reading or auding each passage, examinees must answer 12 questions about each. Questions are of the constructed response type answerable by a word or single phrase. Constructed responses avoid the problem of some correct answers being attributable to guessing, as well as the requirement for producing an adequate number of reasonable distractor alternatives. Answers on both the Auding and Reading Paragraphs subtests must be given in writing. The LAB Paragraphs Test asks questions requiring the recall of facts and avoids those requiring inferential or other higher order reasoning processes. Thus, the Paragraphs Test assesses ability to store factual information presented in spoken or written passages and to retrieve it later in response to a question. While such tasks do not represent the total range of skills involved in reading or listening comprehension, they are important to any sort of learning by language.

2. A Vocabulary Test designed to measure the discrepancy between a person's knowledge of word meanings presented by auding and by reading. The vocabulary words were selected from the Paragraphs Test passages; this was done to determine whether or not poor performance on the Paragraphs Test might reflect lack of knowledge of word meanings. Additionally, LAB Vocabulary scores provide information as to whether or not a person's performance on the Paragraphs Test suffers due to the requirement to process efficiently information in connected prose format. The Vocabulary Test, unlike the Paragraphs Test, imposes no requirement for processing 150 to 190 words of prose and then responding to recall questions, all of which places additional demands on memory and attentional processes. Since the Vocabulary Test is intended to be diagnostic of performance on the Paragraphs Test, it uses important concept words taken from the paragraphs. This is not the situation in typical reading tests.

To construct the LAB Vocabulary Test, fourteen words were selected from each of the auding and reading passages in the Paragraphs Test. Most (93%) of these words had appeared in a retention test item and 85% were found on a basic word list for adults (Mitzel, 1966). In the test, each word is presented in a stem within the context of a short phrase from the appropriate Paragraphs Test passage. Examinees must select the best synonym for this word from four alternatives. The 28 vocabulary words derived from the reading paragraphs are presented for reading. The two sets of words derived from the Auding Paragraph Subtest are presented for simultaneous auding and reading in the Vocabulary test. This latter mode of presentation was chosen to permit the examinees to use auding if reading skills were too low, or reading if so desired. The major aim was to learn if the vocabulary knowledge was available to the examinee, not to assess auding or reading capability in this particular instance. The presentation times for each condition were established and equated by permitting the same amount of time for the reading items as it took to read the simultaneous auding and reading items aloud to examinees a single time.

3. A Decoding Test designed to measure the efficiency with which a reading decoding task can be performed using units of connected discourse. This test represents an attempt to index operationally the degree of "automaticity" of decoding as discussed by Fries (1963). With regard to reading, automaticity refers to the ability to decode print so efficiently that attention can be directed toward the processing of meaning

instead of toward the decoding task. It implies that skill in decoding has become so proficient that decoding can be done pre-attentively, and attention can more effectively be allocated toward conceptualizing the message.

In the LAB Decoding Test, the examinee is required to simultaneously aud and read passages at four different rates of presentation. Rates are established by the spoken message so that the auding presentation rate sets the pace for the reading task. At times, the word being spoken for auding is arranged to differ from the word being read. The examinee's task is to detect and circle the mismatches between the word on the page and the word heard. Since the mismatching words embedded in the passages are semantically and syntactically acceptable, they are not detectable unless the person attends to the spoken presentation at the same time he or she is reading. The four passages used in the test were all selected from the same first aid manual and have a difficulty level of 9th grade across rates. The average length of the four passages used is 332 words, with 11 mismatches per passage. The first of the eleven is considered practice so that a person is scored on ten mismatches per passage. Mismatches occur in each passage at an average rate of one per 30 words. No mismatches are closer than 12 words apart. The four rates, which were selected on the basis of previous work (e.g., Foulke & Sticht, 1969), were 100, 150, 200, and 250 words per minute. These represent a range of ±3 standard deviations from the accepted mean reading aloud rate of 175 wpm.

At the time the LAB was developed (Sticht and Beck, 1976), a small-scale tryout of the LAB was performed using as subjects 78 male inmates at a correctional facility. The mean Gates-MacGinitie reading grade level for this group was 7.8. This study found that the correlation between LAB subtests was quite high, ranging from .53 to .86 for the form which became the current LAB. Correlations between LAB subtests and Gates-MacGinitie scores were also high, from .60 to .86, suggesting that the LAB validly assesses language skills assessed by this widely used standardized reading test.

Because the LAB test assesses auding skill, it may have potential for improving the prediction of the success of lower-aptitude, less literate applicants for military service. Additionally, the LAB may be useful as a diagnostic instrument. When an applicant scores poorly on the written language component, the LAB provides information to aid in deciding whether the person has a reading decoding problem (that is, comprehending language in written form), or is poor in oral comprehension (that is, has a low level of oral language ability) or both. Thus, the LAB is a potentially useful instrument to supplement the ASVAB because it assesses the two major receptive communication modes, auding and reading, and it provides diagnostic information that can be used to identify persons who might benefit from a brief remedial reading program.

NORMING THE LAB

To render the LAB usable as a potential supplement to the ASVAB for screening applicants for military service, research was conducted to develop normative data for the LAB. The remainder of this chapter describes the research to develop normative data to permit the interpretation of LAB scores in terms of percentile scores for the LAB itself, and in terms of reading grade levels for several commercially available, standardized, norm-referenced reading tests. Additionally, the norming study has produced data permitting the interpretation of the LAB and its subtests in terms of Armed Forces Qualification Test (AFQT) percentile scores and in terms of the General Technical (GT) composite scores of the ASVAB. Thus, by administering only the LAB, it is possible to obtain estimates of a person's reading grade level on several tests and AFQT and GT composite scores, in percentiles, on the ASVAB.

The norming study to be discussed in the present chapter also provided data used in the predictive validity study to be discussed in Chapter 4. To avoid duplication, the description of the population sampled in the norming and predictive validity studies will be given in the present chapter. However, only the results for the norming study will be presented in this chapter. Chapter 4 will present the results of the predictive validity study comparing the LAB to the ASVAB as a predictor of successful performance in military jobs.

METHOD

SUBJECTS

The data to be discussed here were obtained from a total population of 4599 service applicants who were tested at 25 geographically dispersed Armed Forces Examining and Entrance Stations (AFEES) in March and April of 1978. Plans called for all applicants to be administered the Armed Services Vocational Aptitude Battery and two of four reading tests. These tests were the Nelson-Denny Reading Test, the Gates-MacGinitie Reading Test, Reading Test of the Basic Skills Assessment Battery and the Literacy Assessment Battery (LAB). These tests are described in the next section. The geographic locations of the AFEES concerned, the number of applicants receiving each test, and the demographic and service relevant characteristics of the population tested are described in Tables 2, 3, and 4.

For the predictive validity study, Defense Manpower Data Center (DMDC) cohort files were searched for the service records of the people in the original AFEES LAB sample (i.e., those who had taken the LAB test). Follow-up service career data were found for 980 of the 2111 members of the original sample. For maximum interpretability of service record data, it was decided to consider only people who had had no prior service and who had either been separated from the service since entry or had served at least 24 months by September 1980. This decision further reduced the number in the sample to 789. Demographic characteristics of this LAB follow-up sample are presented in Table 2. The nature of the criterion variables selected from these service records for use in the predictive validity study will be discussed in Chapter 4.

TESTING PROCEDURE

Every applicant for service who came to the AFEES during a six-week period in March and April 1978 was tested. The ASVAB was administered as part of the standard entrance procedure. Reading tests were administered either on the same day or the preceding evening. The two reading tests given at that station were administered in counterbalanced order by a test control officer. During the first week of testing, reading test administration at each site was overseen by a member of the Joint-Service ASVAB Working Group to make sure that correct administration procedures were followed. Applicants were tested in groups of 10-50.

DESCRIPTION OF TESTS

The Armed Services Vocational Aptitude Battery (ASVAB) used in the present study was the operational selection and classification test for all armed services during the period

Table 2

Demographic Characteristics of LAB Evaluation Population

Characteristic	Total Sample	LAB Sample	LAB Follow-Up Sample
Service			
Army	1930 (44%)	963 (46%)	313 (40%)
Navy	1086 (25%)	497 (23%)	198 (25%)
Air Force	811 (19%)	361 (17%)	150 (19%)
Marine Corps	537 (12%)	285 (14%)	123 (16%)
Qualification			
Qualified	2999 (70%)	1481 (73%)	710 (92%)
Unqualified	1277 (30%)	546 (27%)	58 (8%)
Sex			
Male	3601 (82%)	1741 (83%)	686 (87%)
Female	781 (18%)	365 (17%)	103 (13%)
Race			
Black	1683 (38%)	648 (31%)	205 (26%)
White and Other	2700 (62%)	1444 (69%)	584 (74%)
Education			
High School Graduate	2225 (51%)	1086 (52%)	474 (60%)
Non-Graduate	2110 (49%)	1003 (48%)	316 (40%)

Table 3

Number in AFEES and Follow-Up Samples

Test	Total N	LAB Sample N	LAB Follow-Up Sample N
ASVAB	4260	2111	789
Gates-MacGinitie	2245	593	210
Nelson-Denny	2437	673	299
Basic Skills Assessment	1922	688	244

Table 4

AFEES Stations Where Tests Were Administered

Station	Total Sample N	LAB Sample N
Boston	111	111
Newark	140	-
Philadelphia	502	502
Pittsburgh	95	-
Atlanta	285	-
Jacksonville	35	35
Louisville	205	205
Montgomery	141	-
Raleigh	57	-
Richmond	92	-
Dallas	290	-
Denver	199	199
Houston	200	200
Memphis	161	-
New Orleans	218	-
Oklahoma City	197	-
San Antonio	237	237
Cincinnati	191	-
Detroit	223	223
Indianapolis	224	-
Milwaukee	103	103
Minneapolis	235	235
St. Louis	79	
Salt Lake City	69	69
Fresno	94	94

January 1976 through October 1980. ASVAB Forms 6 and 7 are multiple-choice tests consisting of 13 subtests requiring approximately three hours to administer. A brief description of each subtest is given in Table 5.

Table 5

Subtests of the Armed Services Vocational Aptitude Battery (6/7) Used in This Study

1. General Information (GI): contains questions on factual information which may be considered common knowledge, e.g., what citrus fruits are.
2. Numerical Operations (NO): contains problems of addition, subtraction, multiplication and division with emphasis on speed and accuracy.
3. Attention to Detail (AD): contains items where the examinee must detect and count the number of "c" characters intermixed in rows of "o" characters. Emphasis is on speed, as well as accuracy.
4. Word Knowledge (WK): involves selecting the best synonym for words from among four alternatives.
5. Arithmetic Reasoning (AR): requires the solution of arithmetic word problems and selection of the answer from four alternatives.
6. Space Perception (SP): contains problems requiring the visualization and manipulation of objects in space.
7. Mathematics Knowledge (MK): gives problems in which mathematical principles must be applied.
8. Electronic Information (EI): requires the identification or application of electrical or electronic knowledge.
9. General Science (GS): contains questions tapping knowledge about physical and chemical properties.
10. Mechanical Comprehension: presents an illustration of a mechanical operation and asks a question about it.
11. Shop Information (SI): requires knowledge and familiarity with tools and practices in shop activities.
12. Automotive Information (AI): requires knowledge and familiarity with the maintenance and repair of automotive equipment.

The Armed Forces Qualification Test (AFQT) composite of the ASVAB was, in 1978, made up of scores on the WK, AR, and SP subtests. The AFQT is used by all four services for making personnel selection decisions. The General Technical (GT) composite made up of WK and AR subtests is used for job classification.

The AFQT scores referred to in this paper are given as percentiles, rather than raw scores. These percentiles are referenced to the World War II mobilization population, i.e., all personnel (including officers) on active duty as of December 31, 1944. (OASD(MRA&L), *Aptitude Testing of Recruits*, Report to House Committee on Armed Services, July 1980). All selection and classification decisions made on the basis of

ASVAB scores for the population considered in this study refer to the percentile-raw score equivalences used from January 1976 through September 1980. However, in 1980 the ASVAB Working Group determined that these norms were inaccurate. In particular, the percentile equivalents of lower raw scores were substantially inflated. The most important consequence of this was that a group of people scoring below the median who should have been considered unqualified for service on the basis of their AFQT scores were accepted by the services. When the norming error was discovered, the new correct percentile equivalents to ASVAB raw scores were determined and corrected conversion tables were made available. Unless explicitly stated otherwise, all ASVAB percentile scores given in this paper are based on the new, corrected norms.

In addition to the ASVAB and LAB tests, three commercially available, standardized and norm-referenced reading tests were administered:

The Nelson-Denny Reading Test, Forms C&D (Brown, Nelson & Denny, 1976) is a commercially available instrument designed for use in grades 9 to 12. The Nelson-Denny contains 100 vocabulary items where the closest synonym for a word must be chosen from among five alternatives and 36 multiple choice comprehension questions involving material in eight reading passages. The Nelson-Denny is a speed test in which examinees are allowed 10 minutes for the vocabulary portion and 20 minutes for the comprehension portion. In determining total score, score on the comprehension section is given twice the weight of vocabulary points. The Nelson-Denny was normed on a sample of more than 16,000 high school and college students providing the basis for associating raw scores with percentile ranks and grade level equivalents. Grade level equivalents range between 6.0 and 15.0, with equivalents below 9.0 and above 12.3 having been determined by extrapolation. Split half reliability coefficients for total scores were above .9 for all grades in the norming sample. Test-retest correlations were all above .89.

The Gates-MacGinitie Reading Test, Form D, (Gates and MacGinitie, 1965) is another commonly used commercial test. It has a target population considerably below that of the Nelson-Denny being intended for grades 4 to 6. The Gates-MacGinitie yields two subtest scores, Vocabulary and Comprehension. The Vocabulary subtest consists of 45 items requiring the matching of a word with the correct synonym among five choices. The Comprehension subtest involves very short passages of increasing difficulty in which comprehension is assessed by asking the examinee to select the best alternative for filling two or three omissions in each passage. There are 43 such items. Although a time limit is imposed, the Gates-MacGinitie is a power and not a speed test. Gates-MacGinitie Level D scores can be expressed as reading grade levels ranging from 2.1 to 11.9. Alternate from reliabilities range from .78 to .89.

The Basic Skills Assessment: Reading (Fremer, Swineford and Zieky, 1978) test is one of a battery of three functional literacy tests for secondary and adult students designed to be used to determine if examinees have reached locally determined competency levels. The test developers state that, as the purpose of the test is to discriminate efficiently between individuals requiring basic skills remediation and those who do not, the test is relatively easy for a general population of secondary school students and adults. The Reading test consists of 65 multiple choice items divided between those intended to assess literal comprehension and those assessing "inference and evaluation." In addition, reading questions relate in content to one of five broad functional areas, e.g., consumer activities, and subareas within them. Raw scores on the BSA Reading test are converted to a scaled score for equivalence between test forms and between the Reading, Writing and Mathematics tests of the battery. Scaled scores can be converted to percentiles based on norming groups entering eighth or ninth grade or completing grade 12. No reading grade level equivalences are provided. The Kuder-Richardson split half reliability coefficient of the Reading test is about .94. Summary descriptions of the three commercial tests used are given in Table 6.

Table 6

Description of Reading Tests Used in This Study

Test	Task	No. Items	Time/Rate	Total Score/ Subtest Score	Target Population	RGL Range
Gates-MacGinitie Survey D				Voc + Comp / 2	4th-6th Grade	2.1-11.9
SUBTEST						
Vocabulary	Choose Synonym from 5 alternatives	50	15 Minutes	No. Correct		
Comprehension	Choose from 5 alternatives to fill in blanks in 21 short passages	52	25 Minutes	No. Correct		
Speed-Accuracy	Choose from 4 alternatives to answer questions about 36 short paragraphs	36	5 Minutes	No. Attempted / No. Correct		
Nelson-Denny Form C and D						
SUBTEST						
Vocabulary	Choose closest synonym from 5 alternatives	100	10 Minutes	Voc + Comp / 2	Grades 9-16	6.1-15
Comprehension	Choose from 5 alternatives best answer about 8 paragraphs	36	20 Minutes			
Rate	Number of words read in 1 min.		1 Minute			
Basic Skills Assessment						
SUBTEST						
Reading	Choose best answers to questions in 5 functional domains, e.g., consumer, protection	33 literal comprehension 32 inference and evaluation 65 Total	45 Minutes	No. Correct	Secondary and adult students	Not Used

RESULTS

The presentation of results in this section will be organized as follows. First, the LAB scores obtained by the tested population will be discussed apart from scores on the other tests. Patterns of scores and relationships which appear to hold among LAB component scores will be presented. Next, the relationships between LAB scores and scores of the same population on other tests which were used to norm the LAB will be presented, and then the results of the LAB norming will be given.

LAB SCORES

The means and standard deviations of LAB scores for the population used in this study are given in Table 7.

Table 7

Summary Scores on the Literacy Assessment Battery

LAB Subtest		Total LAB Sample N	Mean	Standard Deviation	Follow-Up Sample N	Mean	Standard Deviation
Paragraphs							
Auding	(out of 24)	2105	14.4 (60%)	5.4	736	14.9 (66%)	4.2
Reading	(out of 24)	2105	15.5 (65%)	6.0	738	17.1 (71%)	4.9
Vocabulary							
Auding/Reading	(out of 28)	2107	24.6 (88%)	4.3	739	25.9 (93%)	2.9
Reading	(out of 28)	2107	23.4 (84%)	5.6	739	25.1 (90%)	3.9
Decoding							
100 WPM	(out of 10)	2107	8.8 (88%)	2.2	739	9.3 (93%)	1.4
150 WPM	(out of 10)	2107	8.2 (82%)	2.6	739	8.8 (88%)	2.0
200 WPM	(out of 10)	2107	6.8 (68%)	3.0	739	7.5 (75%)	2.4
250 WPM	(out of 10)	2107	4.7 (47%)	3.2	739	5.4 (54%)	3.0
TOTAL	(out of 40)	2102	28.5 (71%)	9.7	738	31.0 (78%)	7.4
LAB TOTAL	(out of 144)	2105	106.4 (74%)	27.4	738	115.1 (80%)	19.6

Table 7 shows that performance on the Vocabulary Tests is considerably better in terms of percent correct than is performance on the Paragraphs Test; that Auding/Reading Vocabulary Test scores are higher than Reading Vocabulary Test scores, but that Paragraphs Reading is somewhat better than Paragraphs Auding; that performance at the slowest rate (100 wpm) on the Decoding Test is nearly twice that of the fastest (250 wpm), that these patterns hold for both the LAB total sample and the follow-up sample, and that, overall, the follow-up sample scored higher than the total sample.

To supplement information derivable from mean scores and also as a prerequisite to norming the LAB, frequency distributions of scores were computed for all LAB scores. Figures 3, 4, and 5 present these distributions graphically for LAB component scores.

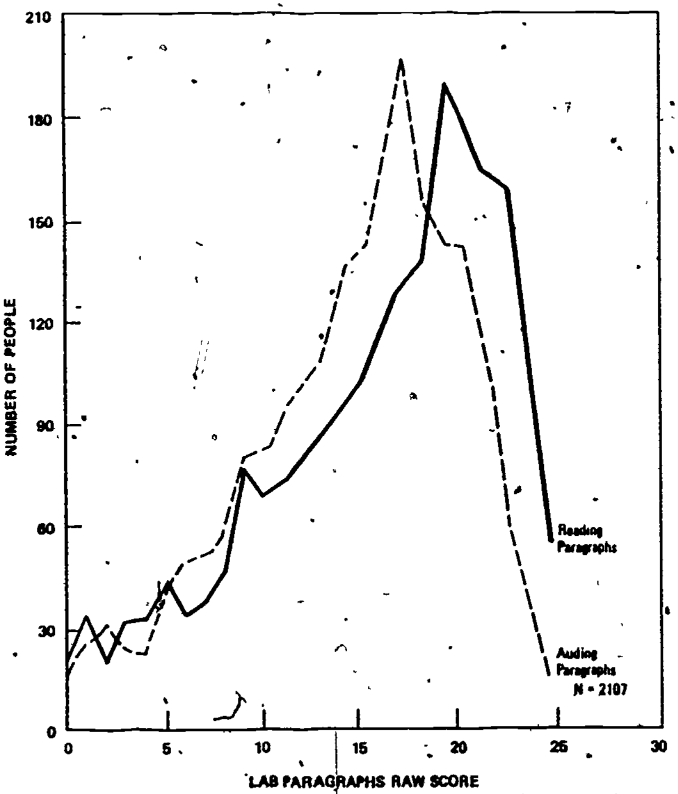


Figure 3. **Frequency Distributions for LAB Paragraph Subtests**

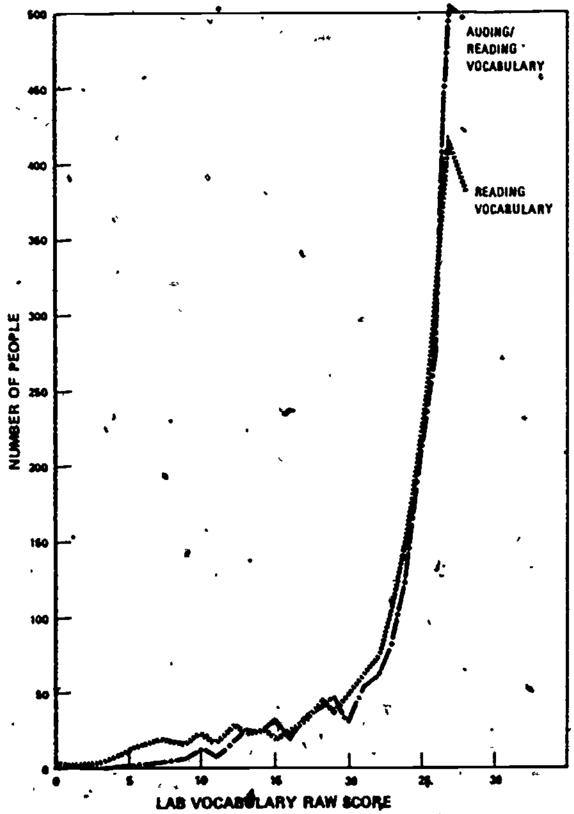


Figure 4. **Frequency Distributions for LAB Vocabulary Subtests**

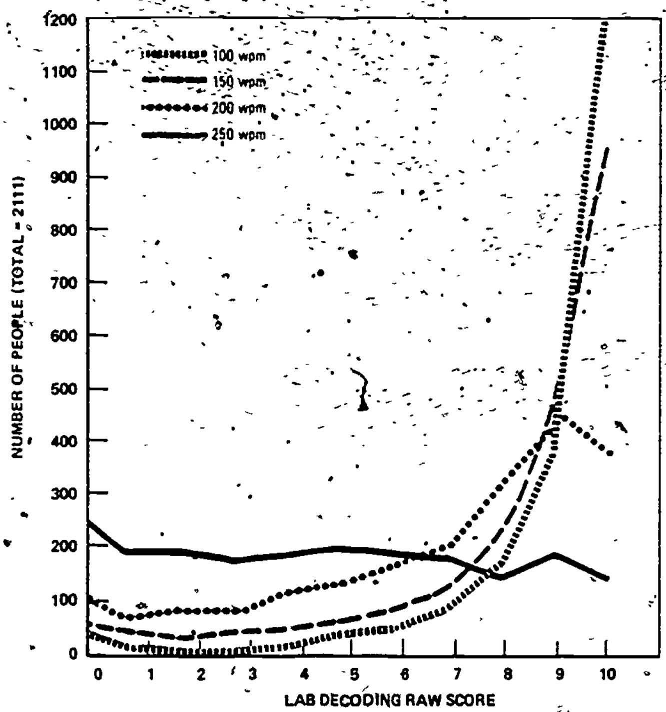


Figure 5. Frequency Distributions for LAB Decoding Subtests

Examination of these frequency distributions reveals that scores on Auding and Reading Paragraph Tests, though somewhat negatively skewed, are essentially normally distributed. Figure 3 also demonstrates graphically that Reading Paragraph performance is superior to Auding Paragraph performance. Figure 4 indicates that the Vocabulary score distributions are highly negatively skewed with more than half the population getting a score of 25 out of 28 or better. Figure 4 shows that scores on Decoding I and II Tests and to some extent on Decoding III are also negatively skewed, indicating that the tests are "easy" for the given population.

Pearson correlations among LAB subtest and total scores are given in Table 8. This table reveals that LAB subtest scores are substantially correlated with each other. Correlations between comparable LAB Auding and Reading scores indicate that people who perform well on one perform well on the other and poor performance scores are also associated with each other. Figures 6 and 7 portray this association. For example, Figure 7 shows that of people who perform in the fourth quartile (lowest 25%) on the Reading Vocabulary subtest, 82% also perform in the fourth quartile on the Auding/Reading Vocabulary subtest, while only 2% perform in the first quartile.

Reliability data for the LAB are given in Table 9. The first set of reliabilities given were computed by means of the Kuder-Richardson 21 formula. These numbers represent an assessment of the internal consistency of these LAB subtests. The Kuder-Richardson 21 computation makes the assumption that all items in a test have the same difficulty. Since this assumption may not hold true for all LAB subtests, an alternate method of estimating reliability was used. In this method, for 100 randomly selected test booklets from the AFEES administration, scores on each of the four Vocabulary and Paragraph subtests were decomposed into two parts, each part being associated with one of the

Table B

Intercorrelations Among LAB Subtests

	LAB Auding Para.	LAB Reading Para.	LAB A/R Vocab.	LAB Reading Vocab.	Decod. 1	Decod. 2	Decod. 3	Decod. 4	Decod. Total	LAB Total
LAB Auding Para.	1.00 (2105)	.749 (2105)	.685 (2105)	.678 (2105)	.481 (2109)	.615 (2109)	.603 (2105)	.579 (2105)	.628 (2100)	.828* (2105)
LAB Reading Para.		1.00 (2105)	.719 (2105)	.762 (2105)	.619 (2109)	.674 (2109)	.728 (2105)	.684 (2105)	.769 (2100)	.906* (2105)
LAB A/R Vocab.			1.00 (2107)	.813 (2107)	.616 (2111)	.627 (2111)	.633 (2107)	.570 (2107)	.689 (2102)	.859* (2105)
LAB Reading Vocab.				1.00 (2107)	.663 (2111)	.705 (2111)	.699 (2107)	.514 (2107)	.754 (2102)	.899* (2105)
Decoding 1					1.00	.810 (2111)	.689 (2111)	.524 (2111)	.828* (2106)	.756* (2109)
Decoding 2						1.00	.792 (2111)	.626 (2111)	.900* (2106)	.811* (2109)
Decoding 3							1.00 (2107)	.779 (2107)	.931* (2102)	.849* (2105)
Decoding 4								1.00 (2107)	.854* (2102)	.784* (2105)
Decoding Total									1.00 (2102)	.906* (2100)
LAB Total										1.00 (2105)

*Indicates correlation of a total score and one of its components.

Numbers in parentheses refer to **N.**

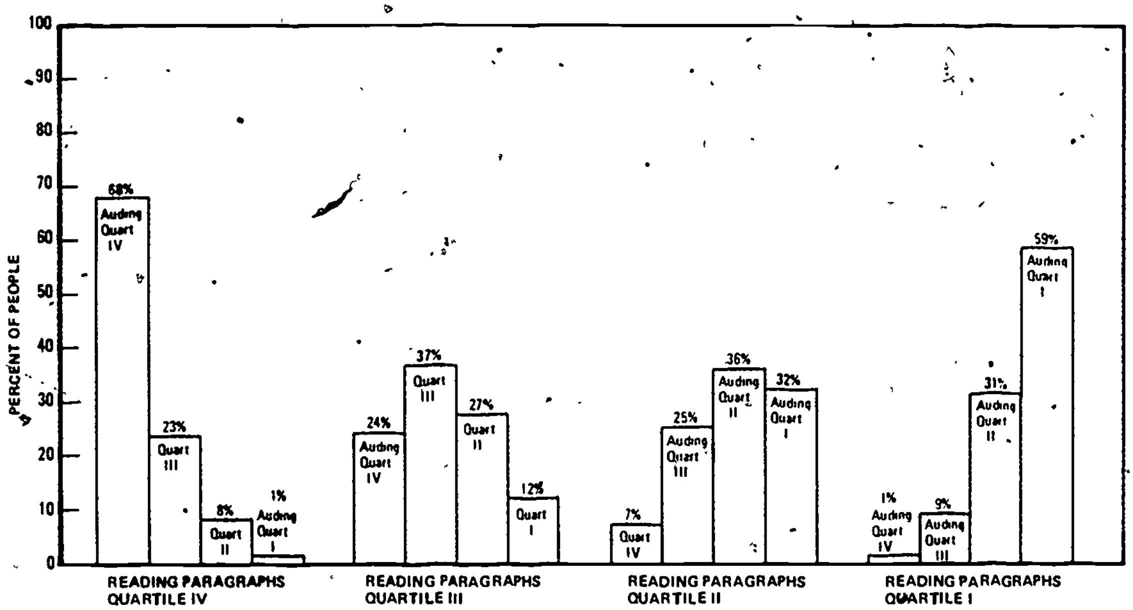


Figure 6. **Distributions of Auding: Paragraphs Quartiles by Reading: Paragraphs Quartiles**

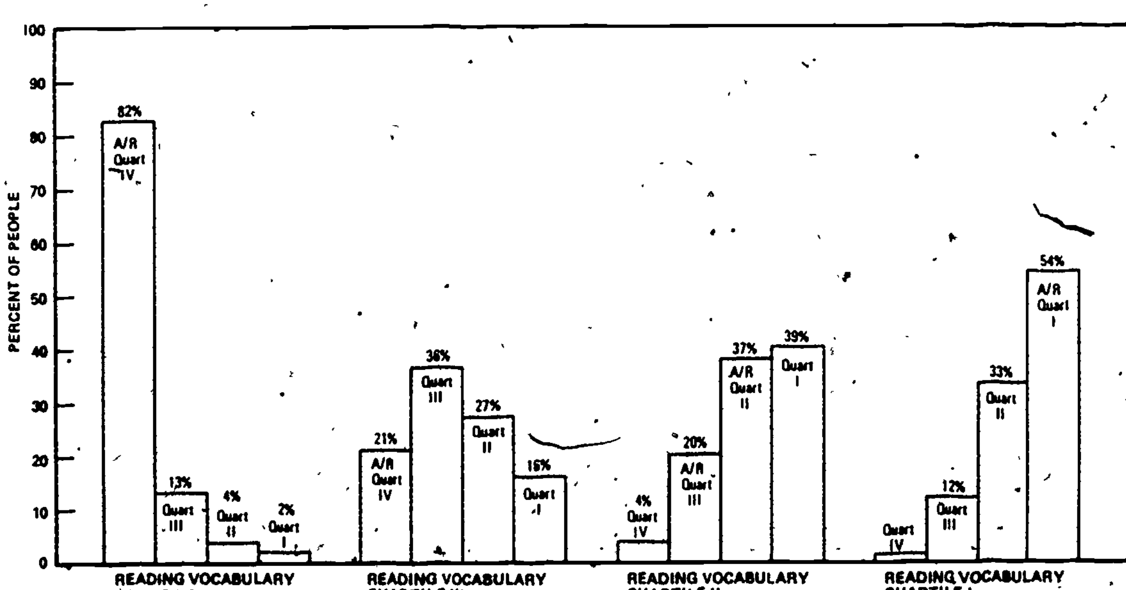


Figure 7. **Distributions of Auding/Reading Vocabulary Quartiles by Reading Vocabulary Quartiles**

Table 9

Reliabilities of LAB Subtests

Test	Number of Items	KR21	r Between Halves
Auding Paragraphs	24	.832	.71
Reading Paragraphs	24	.880	.77
Auding/Reading Paragraphs	28	.875	.79
Reading Vocabulary	28	.912	.80
Decoding I (100 WPM)	10	.869	-
Decoding II (150 WPM)	10	.868	-
Decoding III (200 WPM)	10	.834	-
Decoding IV (250 WPM)	10	.838	-

two passages making up that subtest. Then, a Pearson correlation was computed for each set of 100 pairs of scores. These correlations are also presented in Table 9. It should be noted that because these LAB subtests are timed, these procedures are to be preferred over the split half correlation usually performed to assess reliability (Thorndike, 1971). It can be seen from a comparison of the values of r in Tables 8 and 9, that the correlations between the two halves of the reading and auding subtests are not substantially higher than the correlations between comparable auding and reading subtests.

READING "POTENTIAL" AND CLOSING THE "GAP" BETWEEN AUDING AND READING

As earlier discussion has indicated, the LAB Test provides an assessment of a person's ability to process oral language, in addition to an assessment of reading skills. In particular, the LAB identifies poorer readers whose auding scores indicate "reading potential", the presence of such potential being indicated by auding scores higher than comparable reading scores (see Figure 2). A consideration of Table 7 and Figures 2 and 3 reveals that on the level of group effects there is a tendency for this population to demonstrate "reading potential" in their Vocabulary Test scores, but not in their Paragraphs scores. A frequency count of number of people showing reading potential in LAB Paragraphs and Vocabulary scores indicates that 51% show some vocabulary potential, while 32% show potential in their paragraphs scores. If a difference of three or more points between auding and reading scores is arbitrarily set as an indication of significant potential, the proportion of people with reading potential drops to 21% for Vocabulary and 17% for Paragraphs.

The developmental concept of learning to read as a process of closing the gap between oral and written language comprehension has led to the prediction that more people who are poor readers will show such a gap than people who are good readers (Sticht et al., 1974). One way to verify whether this prediction is confirmed is to look at LAB Reading and Auding scores obtained by people obtaining different reading levels as measured by another reading test. Figure 8 depicts mean LAB Auding and Reading scores obtained by people reading at different grade levels as measured by the Gates-MacGinitie reading test. Figure 9 does the same for people scoring at different deciles of the AFQT.

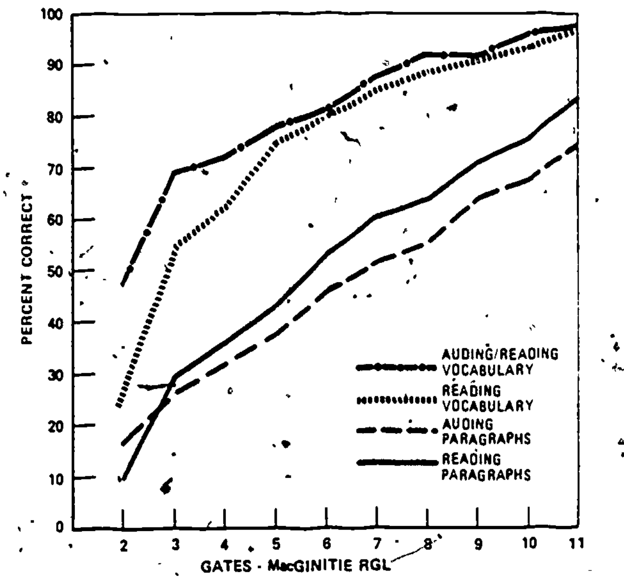


Figure 8. **Mean LAB Paragraph and Vocabulary Subtest Scores as a Function of Gates-MacGinitie Reading Grade Level**

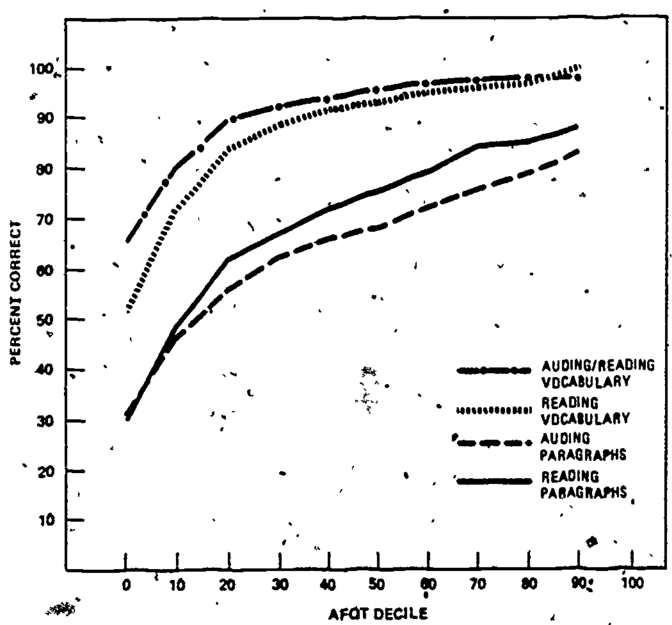


Figure 9. **Mean LAB Paragraph and Vocabulary Scores as a Function of AFQT Decile**

These graphs show that performance on the Auding/Reading Vocabulary test is better than performance on the Reading Vocabulary test for people scoring at all levels of Gates-MacGinitie and AFQT. However, the discrepancy between the two scores decreases as levels of performance on the Gates-MacGinitie or AFQT go up, becoming negligible for the higher levels of performance on these two tests. On the other hand, there is a slight trend, though unreliable with these small numbers, for auding performance to exceed reading performance for paragraphs at the very lowest levels of Gates-MacGinitie and AFQT. For higher levels, Reading scores are higher than Auding, with the difference between the two increasing as performance on the Gates-MacGinitie and AFQT improves.

Another way to look at the relationship between reading skill level and auding is presented in Table 10. This table shows the percentage of people at each level of the

Table 10

Percent of People With Auding Scores Better Than Reading Scores as a Function of Gates-MacGinitie Reading Grade Level and AFQT Decile

Gates-MacGinitie RGL	2	3	4	5	6	7	8	9	10	11+
LAB Vocabulary: N	16	29	35	55	48	50	61	60	65	174
Reading plus auding better than reading (in percent)	88	79	63	51	54	62	57	47	57	38
Reading plus auding more than two points better than reading (in percent)	75	72	46	29	29	28	16	10	5	2
LAB Paragraphs: N	16	29	35	55	48	50	61	60	65	174
Auding better than reading (in percent)	69	41	31	38	31	20	21	28	17	17
Auding more than 2 points better than reading (in percent)	25	31	29	29	15	12	15	10	11	5
AFQT Decile	**0-9**	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90-99**
LAB Vocabulary: N	175	458	219	204	158	218	167	158	187	72
Reading plus auding better than reading (in percent)	71	62	68	58	49	44	39	39	31	17
Reading plus auding more than 2 points better than reading (in percent)	56	39	29	18	16	9	3	3	1	1
LAB Paragraphs: N	175	458	219	204	158	218	167	158	187	72
Auding better than reading (in percent)	44	41	30	32	38	24	28	22	21	15
Auding more than 2 points better than reading (in percent)	29	24	14	18	20	12	13	5	6	3

Gates-MacGinitie or AFQT who scored better by auding than by reading. Clearly, as reading and AFQT scores decline, the proportion of people who score better by auding than by reading increases, indicating that their reading skill is not up to their reading potential.

The LAB Vocabulary Tests measure the knowledge of words taken from the LAB Paragraph Test for the same modality. It is to be expected, therefore, that substantial increases in the LAB Vocabulary scores should be reflected in increases in the LAB Paragraph scores. Figure 10 shows this relationship by plotting the mean Paragraph score for people receiving a given Vocabulary score in the same modality. As earlier figures

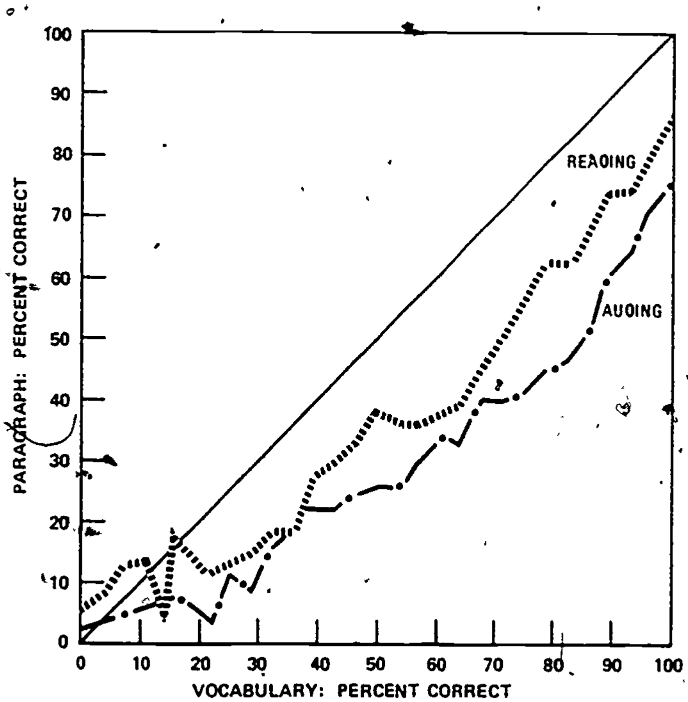


Figure 10. **Mean LAB Paragraph Score as a Function of LAB Vocabulary Score**

have shown, Vocabulary scores are considerably higher than Paragraph scores for virtually the whole range of scores. For example, Figure 10 shows that the people who received Auding/Reading Vocabulary scores of about 80% had a mean Paragraph score of about 45%.

NORMING THE LAB: CONVERSION TABLES

Table 11 gives the mean scores obtained by the total and LAB sample populations on the other reading and ASVAB tests administered. On the basis of the distribution of test scores for the population described above, LAB score conversion tables were prepared by means of an equipercentile norming procedure. These tables, provided in Appendix A, allow the conversion of raw scores on any LAB component or on LAB

Table 11

Mean Scores on Reading Tests and ASVAB Composites for LAB Norming and Follow-Up Samples

Test	Total Sample		LAB Sample		LAB Follow-Up Sample	
	N	Mean	N	Mean	N	Mean
Gates-MacGinitie RGL	2245	8.5	593	8.5	210	9.5
Nelson-Denny RGL	2438	9.8	832	9.7	299	10.0
BSA Standard	1925	145	692	149	244	158
AFQT (Old Norms)	4245	46.1	2167	47.2	767	54.8
AFQT (New Norms)	4245	38.9	2167	34.6	767	49.8
GT Raw Score	4261	28.1	2174	29.2	767	32.9

total score to percentile equivalents, reading grade levels on the Gates-MacGinitie and Nelson-Denny tests, standard scores on the BSA test and percentiles on the AFQT and GT ASVAB composites.

The first conversion for each LAB raw score given in the Tables in Appendix A is percentile equivalent. Percentile equivalents indicate the percentage of people in the norming population scoring at or below a given raw score. Thus, a person receiving a raw score of 17 on the LAB Auding Paragraphs Test (Table A-1) has performed as well as or better than 69% of the norming population. The population used to norm the LAB was composed of applicants for military service at geographically representative AFEES in two months of 1978. Thus, it represents a sample from the manpower pool available to the military at that time. Percentile scores, indicating directly where a person's performance stands relative to this population, should be especially meaningful to military users of the LAB test. Percentile equivalents are also appropriate to making normative comparisons among a person's scores on the various subtests.

The Reading Grade Level (RGL) equivalents of LAB scores were also obtained through an equipercentile norming procedure. Percentile distributions for RGL scores on the standardized reading tests were developed for the test population. Then, reading grade level scores were set equal to LAB scores having the same percentile equivalents.

A comparison of the Nelson-Denny and Gates-MacGinitie columns in the norming tables reveals that the RGL equivalents of the same LAB score can be quite different depending on which reading test is used. For example, a person who scored 8 on the LAB Reading Paragraphs Test has a Gates-MacGinitie RGL equivalent of 4.9 and a Nelson-Denny equivalent of 7.0. These levels appear to imply quite different things about what sorts of tasks this person would be likely to be able to perform. Yet, the question of what reading level this person is "really" reading at is unanswerable because of the differences in the populations the Gates-MacGinitie and Nelson-Denny were intended to assess and were normed with, and the tasks included in the tests. Because of these differences the Nelson-Denny RGL does not reflect discriminations between very low scores on the LAB and the Gates-MacGinitie does not reflect discriminations among very high scores.

REGRESSION NORMS

As an alternative to the equipercentile method for norming the LAB, regression equations were computed for relating LAB scores to Nelson-Denny and Gates-MacGinitie reading grade levels and to AFQT percentiles. These equations are given in Table 12.

Table 12

Regression Equations for Converting LAB Scores to Reading Grade Level on Standard Reading Tests and to AFQT Percentile

Converting LAB to Gates-MacGinitie RGL

RGL =

3.3 +	.33	(Auding Paragraph Score)
2.6 +	.39	(Reading Paragraph Score)
−4.3 +	.52	(Auding/Reading Vocabulary Score)
−2.2 +	.45	(Reading Vocabulary Score)
.69+	.88	(Decoding I Score)
2.3 +	.76	(Decoding II Score)
3.8 +	.72	(Decoding III Score)
5.8 +	.65	(Decoding IV Score)
1.6 +	.25	(Decoding Total Score)
1.6 +	.10	(LAB Total Score)

Converting LAB to Nelson-Denny RGL

RGL =

5.9 +	.27	(Auding Paragraph Score)
5.1 +	.31	(Reading Paragraph Score)
1.9 +	.32	(Auding/Reading Vocabulary Score)
3.3 +	.28	(Reading Vocabulary Score
5.6 +	.48	(Decoding I Score)
5.9 +	.48	(Decoding II Score)
5.9 +	.56	(Decoding III Score)
6.8 +	.61	(Decoding IV Score)
4.8 +	.17	(Decoding Total Score)
2.8 +	.07	(LAB Total Score

Converting LAB to AFQT Percentile

RGL =

− 4.5	+ 3.2	(Auding Paragraph Score)
− 5.0	+ 3.0	(Reading Paragraph Score)
−48.1	+ 3.6	(Auding/Reading Vocabulary Score)
−29.8	+ 3.0	(Reading Vocabulary Score)
− 4.3	+ 5.2	(Decoding I Score)
.89	+ 4.9	(Decoding II Score)
5.1	+ 5.4	(Decoding III Score)
15.9	+ 5.5	(Decoding IV Score)
− 7.5	+ 1.7	(Decoding Total Score)
−32.8	+ .68	(LAB Total Score)

In norming the LAB by the regression method, correlation coefficients were obtained for LAB scores with the other scores available. These intercorrelations are presented in Table 13. Correlations among LAB components are omitted because they appear in Table 8. Table 13 indicates that LAB scores, particularly LAB total scores correlate about as well with scores on other standardized reading tests as these tests do with each other. Although LAB total score has the highest correlation with scores on other reading tests, all the LAB components, including the auding components, correlate substantially with these tests. LAB total score is as highly correlated with the AFQT and GT ASVAB composites as are the other reading tests.

A comparison of Table 8 and 13 indicates that LAB Auding and Reading test scores correlate about as highly with each other as do the two reading subtests within the Nelson-Denny and Gates-MacGinitie tests. These moderately high correlations among LAB scores and scores on other reading tests indicate that members of this population rank themselves quite consistently on the basis of their performance on all these reading tasks. The fact that LAB Auding scores correlate nearly, but not quite as highly with reading test scores as the LAB Reading scores do, is consistent with the hypothesis suggested by the developmental model that a large component of what reading tests measure is languaging ability. This ability is not specific to reading alone, but common to performance with both written and oral language.

The substantial correlations between scores on the ASVAB composites and measures of reading might be interpreted to indicate that the ASVAB does a reasonably good job of assessing reading and languaging skills. This is hardly surprising since one component of the AFQT and GT composites is the Word Knowledge (WK) test which is no different in kind from the vocabulary components of the reading tests used. The other two ASVAB tests contained in the AFQT (AR;SP) would on the surface seem to be assessing something other than reading. However, decoding and understanding each word problem is a necessary prerequisite for correct responses on the Arithmetic Reasoning (AR) test. Even the Spatial Perception test requires that rather complex task instructions be read and understood. The understanding of these instructions and use of them to induce a procedure for performing the spatial perception tasks correctly may pose a more demanding application of language skills and knowledge than the selection of the appropriate synonym for a vocabulary word.

THEORETICAL INTERPRETATION OF INTERCORRELATIONS AMONG LAB AND OTHER TESTS

The relationships among the oracy and literacy tests shown in Table 13 can be further understood in terms of the developmental model. According to that model, the prelinguistic child of Stage 2 (Figure 1) represents knowledge in non-language forms and picks up new information through listening and looking (excluding other processes such as touching, etc. for the present purposes). This mode of representation is primarily perceptual, rather than conceptual and has been labeled by some as "imagery" or "analogic" or "spatial" (Marshalek, 1981).

According to the developmental model, the imagery or spatial (auditory and visual) adaptive processes develop first, and then they transform to form oral language (listening changes to combine listening and languaging and is called auding), and the latter in turn is integrated with looking skills to produce looking and languaging which is called reading. According to this way of thinking, all reading test performance includes reading, languaging, and imagery. All auding test performance includes languaging and imagery. Tests such as the Space Perception subtest of the ASVAB can be taken as indicators of imagery. Of

Table 13

Intercorrelation Among LAB, Reading Test, and ASVAB Subtests

	Gates-MacGinitie Vocabulary	Gates-MacGinitie Comprehension	Gates-MacGinitie Total	Nelson-Denny Vocabulary	Nelson-Denny Comprehension	Nelson-Denny Total	Basic Skills Assessment	ASVAB Word Knowledge	ASVAB Arithmetic Reasoning	ASVAB Spatial Perception	ASVAB Percent	GT Raw
LAB Auding Paragraph	.70	.71	.74	.52	.53	.56	.62	.61	.57	.34	.64	.67
LAB Reading Paragraph	.78	.76	.80	.62	.66	.68	.76	.66	.62	.34	.66	.71
LAB A/R Vocabulary	.72	.70	.74	.50	.50	.53	.66	.61	.53	.34	.58	.63
LAB Reading Vocabulary	.80	.73	.80	.55	.57	.59	.69	.66	.55	.35	.62	.68
LAB Decoding Total	.77	.73	.78	.57	.61	.63	.75	.65	.57	.31	.62	.67
LAB Total	.86	.82	.88	.63	.66	.69	.81	.73	.65	.38	.71	.76
G-M Vocabulary	1.00	.80	.94*	.71	.64	.70	.79	.77	.59	.35	.72	.77
G-M Comprehension	.80	1.00	.96*	.61	.58	.62	.81	.70	.61	.41	.71	.73
G-M Total	.94*	.96*	1.00	.69	.64	.70	.84	.77	.63	.40	.75	.79
N-D Vocabulary	.71	.61	.69	1.00	.78	.94*	.63	.74	.58	.31	.72	.74
N-D Comprehension	.64	.58	.64	.78	1.00	.95*	.68	.65	.55	.29	.66	.67
N-D Total	.70	.62	.70	.94*	.95*	1.00	.72	.74	.60	.32	.73	.75
BSA	.76	.81	.84	.63	.68	.72	1.00	.71	.64	.41	.70	.74
ASVAB Word Knowledge	.77	.70	.77	.74	.65	.74	.71	1.00	.65	.38	.88*	.86*
ASVAB Arithmetic Reasoning	.59	.61	.63	.58	.55	.60	.64	.65	1.00	.46	.85*	.86*
ASVAB Spatial Perception	.35	.41	.40	.31	.29	.32	.41	.38	.46	1.00	.67*	.45
AFQT Percent	.72	.71	.75	.72	.66	.73	.70	.88*	.85*	.67*	1.00	.95*
GT Raw	.77	.73	.79	.74	.67	.75	.74	.95*	.86*	.45	.95*	1.00

*Indicates correlation of a total-score and one of its components.

course, there are also task factors, such as the reasoning and special knowledge demands that influence actual test performance using literacy, oracy, or imagery tests. Further, the actual assessment of skills in tests such as those of Space Perception generally involve the ability to comprehend oral and written language to follow the test instructions. But performance on such tests is not as much dependent on language as is performance on oracy and literacy tests.

Given the developmental sequence: imagery → oracy → literacy, and using as indicators of these skills the Space Perception test of the ASVAB (imagery), the LAB Auding Paragraphs (oracy), and the LAB Reading Paragraphs (literacy), we can expect to find the Space Perception test least correlated with literacy tests, the LAB Auding Paragraphs test correlated more highly with literacy tests, and the LAB Reading Paragraphs test the most highly correlated with literacy tests.

Table 14 presents correlations of the ASVAB Space Perception test, LAB Auding and Reading Paragraph tests to each other, to education level, and to the various subtests of the standardized reading tests, and the subtests and composites of the ASVAB.

The first column of Table 14 shows that Space Perception correlates .34 with both the LAB Auding and Reading Paragraphs tests, indicating that, as suggested by the developmental model, imagery underlies and is involved in both oral and written language processing. The second column of Table 14 shows that the Auding and Reading Paragraphs subtests are fairly highly correlated (.73) as would be expected from the developmental model, the similarity of the attentional and semantic memory task demands, and the close temporal proximity in assessment with the two subtests.

Education is more highly correlated with the Auding and Reading Paragraphs tests than with the Space Perception test as might be expected since one function of schooling is to systematically develop language skills. The fairly low correlations of education with Auding and Reading no doubt reflect the restricted range of education levels, and the extreme over-representation (over 50%) of people in the category of "12 years or more" or high school completion.

The Attention to Detail (AD) subtest of the ASVAB could probably serve as another proxy for pre-linguistic information processing abilities much as the Space Perception test is used in the present analysis. The slightly higher correlation of AD to the Reading Paragraphs test may reflect the fact that alphabetic letters are used in the AD task, and hence some slight relationship to decoding might be influencing the AD and LAB Paragraph Reading correlation.

Beyond the first four columns of Table 14, the relationships of LAB Decoding Test total scores and the various vocabulary and paragraph tests to the Space Perception, LAB Auding Paragraphs and Reading Paragraphs test present a recurrent pattern that is consistent with the developmental sequence outlined above: imagery → oracy → literacy. Across the eight columns, the data show that Space Perception has the lowest correlations, Auding Paragraphs the next highest correlations, and LAB Reading Paragraphs the highest correlations with the various literacy tests. The same pattern is found for the Space, Auding, and Reading row tests with the various literacy test total scores and the GT composite from the ASVAB (far right-hand columns of Table 14). On the average, the ASVAB Space Perception test correlates about .35, the LAB Auding Paragraphs about .63 and the LAB Reading Paragraphs about .73 with the various literacy tests.

The effects of <u>special knowledge</u> on the pattern of correlations is indicated by the correlations between the three row tests and the ASVAB special knowledge tests. For six of the nine special knowledge tests, the pattern of correlation shows Space Perception with the lowest, LAB Auding Paragraphs the next highest, and the LAB Reading Paragraphs with the highest correlations with special knowledge, preserving the pattern found

Table 14

LAB and ASVAB Intercorrelation Matrix Interpreted in Accordance With the Developmental Model of Literacy

	Space Perceptions	Auding Paragraphs	Education	Attention to Detail	LAB Decode Total	LAB Vocab A&R	LAB Vocab R	G-M Vocab	ASVAB WK	N-D Vocab	G-M Paragraphs	N-D Paragraphs	Math-Knowledge	Num. Operations	Arith. Reasoning	Electronics (EI)	General (GI)	Science (GS)	Automotive (AI)	Shop (SI)	Mech. Comp. (MC)	LAB Total	G-M Total	N-D Total	BSA Total	GT	AFQT
						Vocabulary Tests					Paragraph Tests		ASVAB Special Knowledge Tests: Mathematics			Other Knowledge						Composite Test Scores: Literacy Tests				ASVAB	
Space Perception (means)						(x̄=.35)					(x̄=.35)		(x̄=.43)									(x̄=.39)					
Space Perception	1.00	.34	.06	.22	.31	.34	.35	.35	.38	.31	.41	.29	.41	.31	.46	.47	.38	.42	.42	.47	.56	.38	.40	.32	.41	.45	.64*
Auding Paragraphs (means)						(x̄=.63)					(x̄=.62)		(x̄=.53)									(x̄=.64)					
Auding Paragraphs	.34	1.00	.20	.20	.61	.67	.66	.70	.61	.53	.71	.53	.53	.47	.57	.57	.57	.58	.48	.51	.50	.82*	.74	.56	.62	.65	.63
Reading Paragraphs (means)						(x̄=.72)					(x̄=.71)		(x̄=.56)									(x̄=.73)					
Reading Paragraphs	.34	.73	.23	.30	.77	.72	.76	.78	.66	.62	.75	.66	.61	.59	.63	.60	.58	.63	.43	.49	.51	.91*	.80	.68	.77	.71	.68

*Part-whole correlations.

between the row tests and the various literacy tests. However, while the pattern is the same, the magnitude of the correlations is different, with Space Perception tending to have higher correlations, while the LAB Auding and Reading Paragraph correlations tend to shrink.

The latter findings are consistent with the developmental model in showing the effects of essentially two major kinds of processing skills, those that construct internal language from external language displays such as speech and writing, and those prelinguistic in origin, that result in the construction of internal, non-linguistic images from external information displays and task requirements. Among the types of tasks that result in the construction of internal images or spatial modes of thought are mathematics tasks. Concepts such as "place value" are essentially spatial. Mathematics tests require both language and spatial perceptual processing (Marshalek, 1981).

The Mechanical Comprehension (MC) subtest of the ASVAB shows the influence of space perception most dramatically. This test requires the examinee to mentally image and rotate gear assemblies to determine what the consequences would be of turning a gear or pushing a lever in a certain direction. As Table 14 shows, the Space Perception test correlates slightly higher with MC than do the LAB Auding and Reading Paragraphs tests.

More will be said in Chapter 5 about the implications for aptitude assessment of the patterns of intercorrelations in Table 14.

Chapter 4

PREDICTIVE VALIDITY OF THE LAB AND ASVAB

As mentioned in Chapter 3, earlier research sponsored by OASD(MRA&L) suggested that assessment of oracy skills in addition to the literacy skills assessed in the AFQT might permit a more accurate identification of less literate personnel who could, by virtue of their oral language skills, benefit from military training and perform satisfactorily on the job (Sticht, Kern, Caylor, and Fox, 1970).

In this chapter we report research to evaluate this hypothesis using the Literacy Assessment Battery (LAB). Because the LAB assesses both oracy (auding) and literacy (reading) skills, it permits the evaluation of the foregoing hypothesis.

METHOD

In the present research, LAB and ASVAB scores were used to predict several criteria of job performance. The samples of personnel for whom predictor and criterion relationships were obtained were drawn from the test population described in Chapter 3, and differed depending upon the particular criterion variable being predicted. Details of these various samples are discussed in the next section which describes the criterion variables used.

CRITERION VARIABLES

Of the 2,111 people who were given the LAB in 1978, records were available for 551 with no prior service who were still in the service by September 1980. These people may be called the survival sample. Decisions to avoid confounding variables by excluding service members with prior service and non-attritees who had served fewer than 24 months because of delayed entry eliminated 258 people from consideration. Undoubtedly, records of some of the people in the AFEES sample could not be located on the follow-up tape because of irregular or missing data.

However, a significant number of people in the AFEES sample who were not in our survival sample were lost to the military either because they were not allowed to enter or because they entered and then dropped out. Both people who would have succeeded in the military, but did not enter and those who were accepted and dropped out represent losses. It is clearly to the military's benefit to be able to identify people who are unlikely to survive as early as possible and reject them, while not denying enlistment to those who have a good chance to succeed.

Performance on cognitive tests such as the ASVAB and LAB have long been used as predictors of success in the military. For this reason, attention in this study has focussed on using the LAB for predicting military survival or its opposition, attrition.

The selection of criterion variables for evaluating the predictive validity of the LAB test was constrained by the information on data files from original AFEES testing and the information on tested individuals available in Defense Manpower Data Center (DMDC) cohort files. In addition to the necessarily limited number of variables recorded in these files, available data was limited by the fact that AFEES testing was performed in March and April 1978 and the last available update of DMDC's files at the time these analyses were carried out was in September 1980. Thus, records were available, at the most, only for the first 30 months of an individual's service career. This precluded use of variables such as eligibility for retention, long term promotion records, etc.

Under these constraints, the variables which were selected as criteria for the predictive validity portion of this study were: qualification status, probability of being separated from the service for failure to meet minimum behavioral or performance criteria, probability of such separation occurring within the first six months of service or subsequently, months of service before separation, and highest paygrade achieved.

The relationships among these variables and between them and the predictors for the population considered are presented in the flow chart of Figure 11, which traces the progress through various decision points in the military careers of the sample of people receiving the LAB test in AFEES stations in 1978. The criterion variables used are described in Table 15.

The first point where people are lost to the military is at the decision to award qualification status to an applicant. This point corresponds to the first decision point in Figure 11. As can be seen in that figure, in our sample 1481 (70%) people were considered qualified, 546 (26%) were not qualified, and data on qualification were missing for 84 (4%). Although some non-qualified people do obtain waivers and enter, the proportion of accessions is much smaller than for qualified personnel. Qualification status decisions are based on an individual's education and AFQT performance. These standards vary between services and they can change from year to year. Since in 1978 these standards were based on norms which have recently been corrected (Sellman and Valentine, 1981), it has seemed of interest to investigate how well current ASVAB scores and LAB could predict qualification. This was the first analysis performed in the predictive validity study. Further information on the qualification status variable can be found in Table 15.

The next group of criterion variables used involved attrition. People who attrite from the military for one or another reason after they have entered represent a significant loss in money and other resources to the services. Of our initial sample, a total of 238 people of the entering 789 were lost due to attrition. As Figure 11 shows, of these 123 were lost in the first six months and 115 subsequently.

When considering the criterion variable of attrition, a decision was made to look not only at total attrition but also to divide the period for which records were available into two parts, the first six months of service and months 7-30, and to analyze data separately for each period. This division reflects the fact that training and orientation constitute a major portion of the recruits' activities during the first six months in all services and career fields, while first-term service after the first six months is mainly devoted to job performance. In this regard, it should be noted that all but three of the trainee discharges for which records were available occurred during the first six months.

A major reason for analyzing attrition during the first six months and for months 7-30 separately is that literacy and other cognitive demands are quite different in training and for operational service, with training demanding more reading to learn than is required by operational job performance (Sticht, Fox, Hauke, and Zapf, 1976). Thus, causes for attrition are likely to differ for these periods. For example, Sacher and Duffy (1977) found that reading level was more closely associated with attrition early in Navy service than with later attrition. The predictive power of a language test like the LAB might be expected to be greater for attrition during initial entry training than for subsequent periods of service when learning by language demands are less stringent.

Table 15

Criterion Variables Used in Predictive Validity Study

Qualification Status indicates whether an applicant at an AFEES station was rated as qualified or unqualified for accession into a service. Qualification status is thus a dichotomous variable with values of 0 and 1. Of the 2,111 who were tested at the AFEES station, 118 were excluded from this analysis because of missing data. Of the 1,993 remaining, 1,463 (73%) were rated as qualified and 530 (27%) as unqualified.

Attrition for Failure to Meet Minimum Behavioral or Performance Criteria is a dichotomous variable indicating whether an individual who entered a service, attrited subsequently for one of the reasons listed in Table 16 or did not attrite during the period in question. (People who attrited for other than performance or behavioral reasons were not included in the calculation of predictive validity coefficients in either the attrition or non-attrition group, (see Table 16).

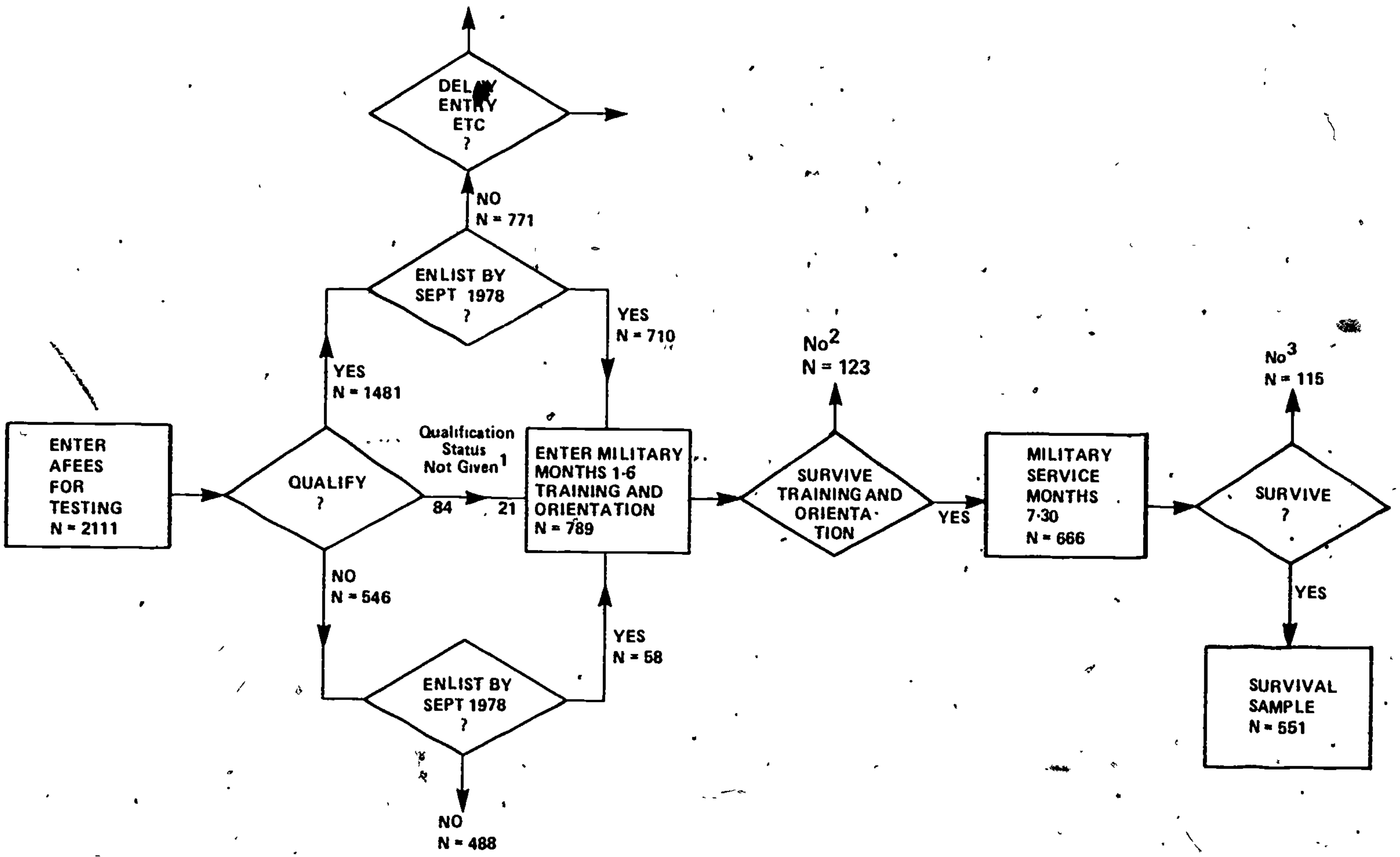
Attrition: 1-6 Months is a dichotomous variable indicating whether an individual who had entered the service attrited for performance or behavioral reasons within six months of enlistment or was still serving at the end of that period. Of 789 accessions, 79 were excluded from the analysis because of missing data and an additional 46 who had attrited in the first six months for other than performance related reasons were also excluded. Of the remaining 664, 67 (10%) attrited in their first six months of service, and 597 (90%) were still in the service after six months.

Attrition: 7-30 Months is a dichotomous variable indicating whether an individual who had served at least six months had attrited by September 1980 or was still serving by that time. Of the 666 people who were still in the service seven months after entering, 64 were excluded from analysis because of missing data and an additional 42 who attrited for other than performance related reasons, were also excluded. Of the remaining 560, 68 (12%) attrited in months 7-30, and 492 (88%) were still in the service in September 1980.

Attrition: Total is a dichotomous variable indicating whether an individual attrited from the service for reasons indicated in Table 16 during the 30 months, for which we have records or was still in the service by September 1980. Of the 789 people who entered the service, 74 were excluded from the analysis because of missing data and an additional 88 who attrited from the service for other than performance related reasons were also excluded. Of the remaining 627, 135 (22%) attrited for performance or behavioral reasons some time in the 30 months, while 492 (78%) were still in the service in September 1980.

Months Before Separation is a continuous variable indicating, for those who attrited from the service for behavioral or performance reasons, how long an individual had served before separation took place. The values of this variable ranged from 1 to 29 months, with a mean of 9.3. Of the 150 people who attrited for performance related or behavioral reasons, 15 were excluded from analysis because of missing data.

Paygrade is a continuous variable indicating for all accessions, the highest paygrade a person had achieved before separation or by September 1980. The values of this variable ranged between 1 and 5 with a mean of 2.9. Of the 789 accessions, 85 were eliminated from this analysis because of missing data.



[1] Data on qualification status was missing for 84 of the people tested, of these 21 were recorded as entering the military.

[2] Includes 77 cases of attrition for failure to meet performance and behavioral criteria and 46 cases of attrition for other reasons (cf. Table 15).

[3] Includes 73 cases of attrition for failure to meet performance and behavioral criteria and 42 cases of attrition for other reasons.

Figure 11. Flow-Chart of Attritions of LAB Sample From AFEES Testing to 24-30 Months After Accession

Because the LAB and ASVAB are cognitive tests, prediction of attrition has been limited to attrition for reasons involving behavior and performance. Thus, we eliminated from consideration attrition for reasons of medical disability, hardship, etc. A complete list of the reasons for attrition included and excluded from the predictive validity analysis can be found in Table 16. Of the 123 people, who, as Figure 11

Table 16

Reasons for Attrition (Interservice Separation Codes) During and Subsequent to First Six Months of Service

Interservice Separation Code	Reason	No. of Attrites During Months 1-6	No. of Attrites After Month 6
FAILURE TO MEET MINIMUM BEHAVIORAL OR PERFORMANCE CRITERIA			
60	Character or Behavior Disorder	3	8
61	Motivational Problems	4	5
64	Alcoholism	0	1
65	Discreditable Incidents	3	8
67	Drugs	0	7
74	Fraudulent Entry	6	3
75	AWOL, Desertion	0	1
76	Homosexuality	1	2
78	Good of Service	3	7
86	Expeditious Discharge	21	28
87	Trainee Discharge	36	3
TOTAL		77	73
OTHER REASONS			
1	Term Expiration	0	1
8	Early Release	1	1
10	Medical Conditions Prior to Service	5	3
11	Disability	0	8
13	Temporary Disability	0	4
16	Unqualified for Active Duty	15	1
22	Dependency or Hardship	1	5
32	Non-Battle Death	0	1
40	Officer Commissions	0	3
91	Erroneous Enlistment	18	0
94	Pregnancy	0	11
97	Parenthood	0	1
98	Breach of Contract	6	0
99	Other	0	3
TOTAL		46	42

indicates, attrited during the first six months, 10 had test scores missing and were eliminated from further analysis, and 67 attrited for reasons of performance or behavior and were thus included as attritees in the predictive validity calculation. Of the 115 attriting after six months, 73 were included in this calculation. The remaining 88 attritees, whose attrition was attributed to other than performance or behavioral reasons, were not considered in the calculation. Further information about these criterion variables is available in Table 15.

The only other available criterion variables indicative of military success identified in the DMDC cohort tape were months of military service before separation and paygrade achieved. Since all the members of the sample who were still on active duty were in the middle of their first enlistment term, total months of military service serves as an indicator of relative career success only for people who had attrited from the service. Thus, the number in the subsample for testing the validity of the LAB in predicting this variable was quite small. Although the other criterion, paygrade achieved, is an appropriate indicator of career success for all members of the sample, its range of values is limited by service requirements for minimum and maximum time in grade. Further information about these criterion variables is available in Table 15.

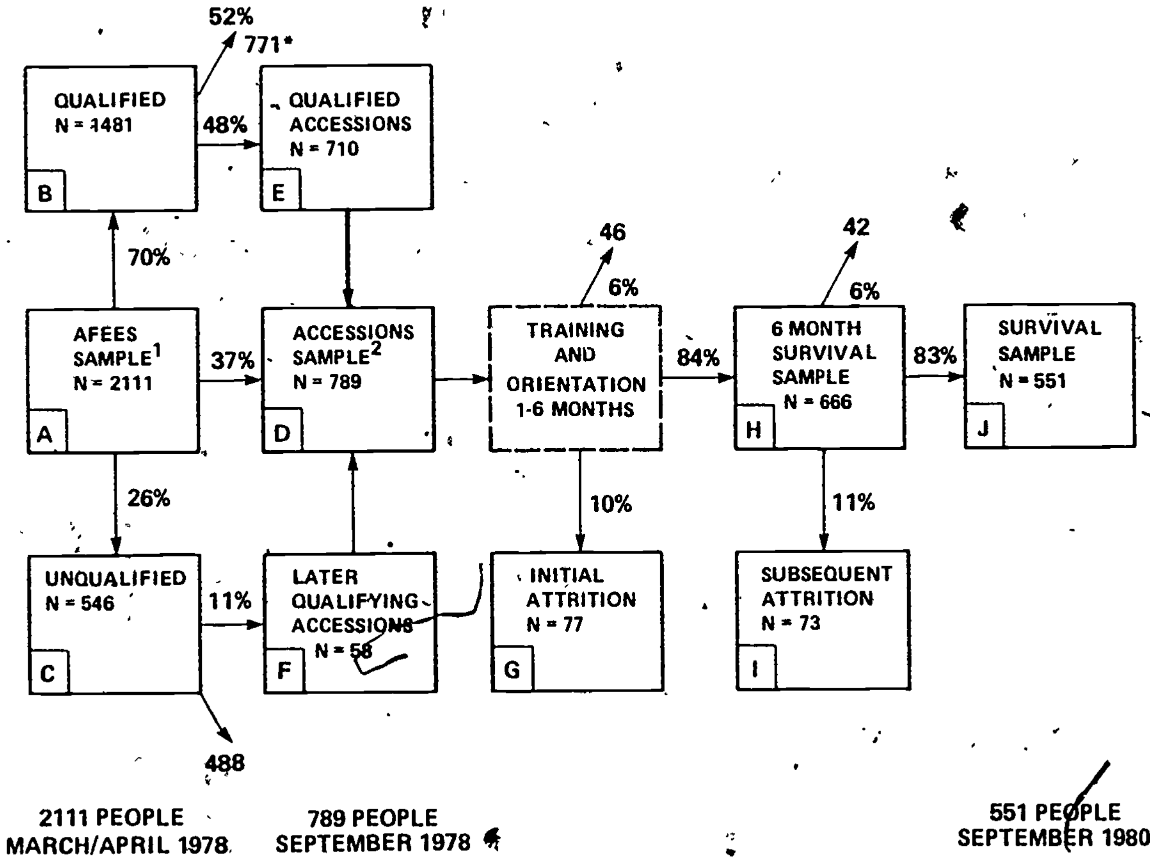
RESULTS

Figure 12, which corresponds to Figure 11 in the previous section, depicts the relationships among the various subsamples considered in the predictive validity portion of this study. Each subpopulation is identified by a capital letter. Table 17 presents mean test scores and demographic and criterion variables for each of these populations.

Table 17 shows that the qualified and unqualified segments of both the initial AFEES population and the accessions population differ sharply in their scores on the ASVAB and other reading tests. This of course, is a consequence of the fact that AFQT is one of the major factors on which qualification decisions are made and the other test scores correlate highly with AFQT and GT. The fact that the scores of initially unqualified accessions are somewhat higher than those of the unqualified population in general is consistent with the idea that the unqualified people who eventually do obtain a waiver and enter military service, come from the more able portions of the unqualified group.

In accordance with expectations, the group which attrites during the first six months of service has poorer test scores than the group which attrites subsequently, and both perform less well than the group that remains. However, these differences are not nearly as striking as between the qualified and nonqualified groups. This undoubtedly reflects the fact that cognitive ability is only one of the factors influencing attrition. It is also evidence that selection decisions awarding qualification status and waivers are successful in screening out those people whose cognitive and language abilities are too low for military success.

The successive elimination of subpopulations with test scores below the mean for the whole population leads to a gradual increase in the mean scores of the subpopulation remaining in the services. This is depicted graphically in Figure 13. It will be noticed in Figure 13 that the major increase in scores takes place between the AFEES sample and the accessions sample. This reflects the fact that the major cognitive screening of the military population takes place before accession.



[1] Includes 84 with no qualification status.

[2] Includes 21 with no qualification status

*This portion of the AFEES sample includes people whose records were missing necessary information on either or both data tapes, people whose entry into service was delayed past September, 1978, people with prior service as well as genuine non-accessions.

Figure 12. **Rates of Attrition of LAB Sample From AFEES Testing to 24 - 30 Months After Accession (Approximate Percents)**

Table 17

Description of Populations Involved in Predictive Validity Study

	AFEES Sample A	Qualified B	Unqualified C	Accession Sample D	Qualified Accessions E	Later Qualifying Accessions F	1-6 Month Attrition[1] G	6 Month Survival Sample H	6-30 Month Attrition[1] Sample I	Survival Sample J
N	2111	1481	546	789	710	58	77	666	73	55[illegible]
Percent High School Graduates	51	56	38	60	59	60	42	62	51	63
Percent Female	18	16	23	13	12	19	17	12	11	10
Percent Black	39	30	61	26	24	53	28	27	25	28
AFQT Percentile	38.9	49.7	13.4	49.8	52.3	17.7	41	50.9	44.8	51
GT	37	51	13	51	51	17	37	51	45	51
LAB Auding Paragraph[2]	14.4(47)	16.1(60)	9.9(22)	15.9(58)	16.1(60)	13(38)	13.7(43)	16.1(60)	14.7(49)	16.3(62)
LAB Reading Paragraph	15.5(45)	17.4(56)	10.3(22)	17.1(55)	17.4(56)	12.5(31)	14.4(39)	17.4(58)	16.3(50)	17.4(56)
LAB Total	106(42)	116(57)	80(15)	115(49)	117(57)	95(25)	104(35)	116(57)	113(42)	116(59)
G-M RGL	8.5	9.4	6.1	9.5	9.7	7.1	8.2	9.6	9.3	9.6
N-D RGL	9.8	10.3	8.7	10.0	10.1	8.4	8.8	10.2	9.6	10.2
BSA Standard	145	156	124	159	160	133	143	160	159	160
Percent Qualified	70	--	--	92	--	--	92	93	89	93
Percent Accessions[3]	37 (50)	48 (62)	11 (20)	--	--	--	--	--	--	--
Percent Attrition[1]	--	--	--	19	19	24	--	11	--	--
Mean Months of Service	--	--	--	21.7	21.8	20.3	2.5	25.3	16.6	27.2
Mean Paygrade[4]	--	--	--	2.9	2.9	2.6	1.2	3.2	1.9	3.4

[1] Includes only attrition attributable to failure to meet behavioral or performance criteria. (cf. Table 16)

[2] Numbers in parenthesis after LAB raw scores refer to percentile equivalents based on the entire norming population.

[3] Does not include individuals with prior service nor delayed entry accessions. Numbers in parentheses indicates accession rate for all those in sample we have records for.

[4] Paygrade scale includes values of E1, E2, E3, E4, E5, E6, E8, E9.

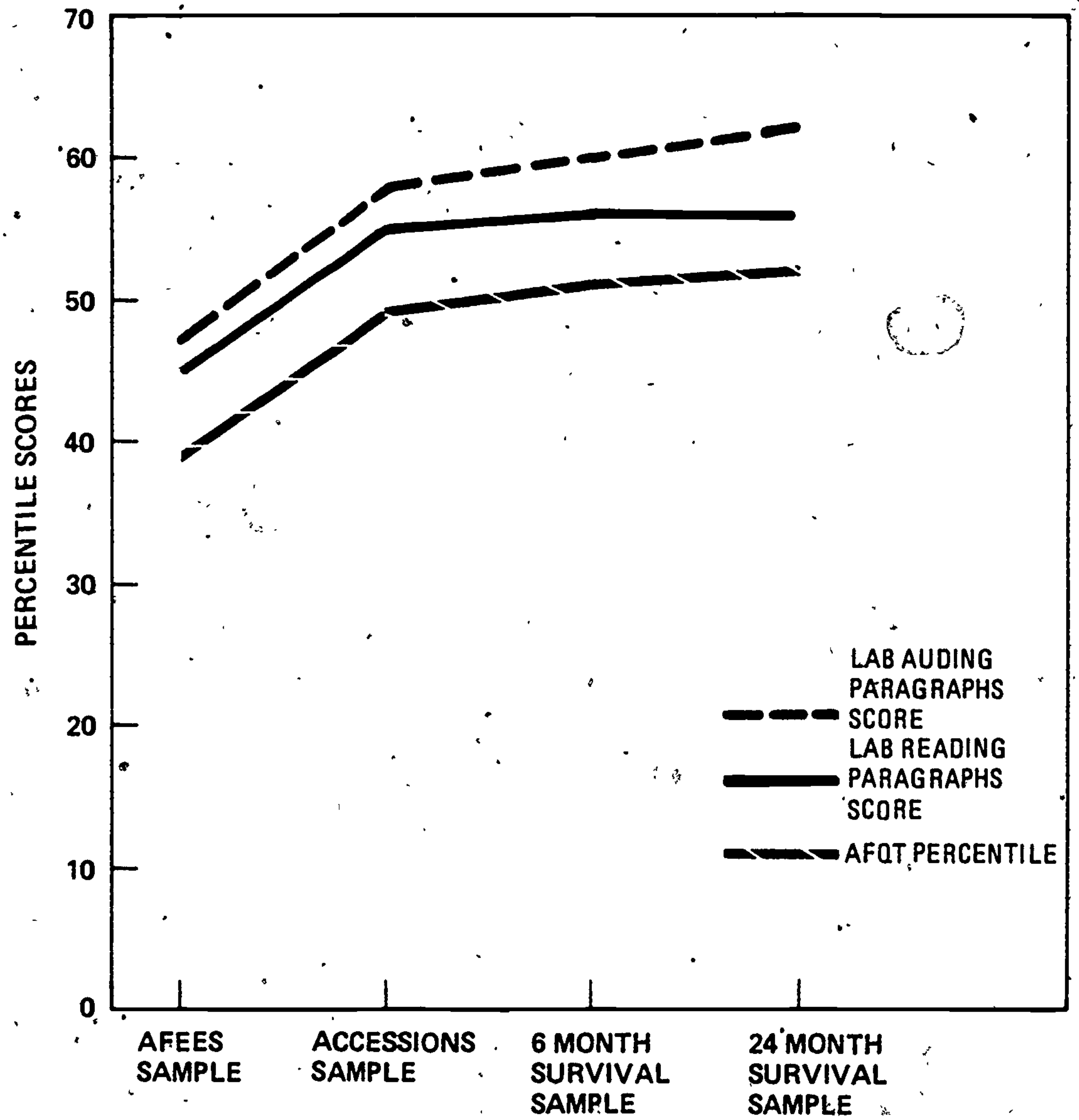


Figure 13. **Changes in Test Scores as a Function of Duration of Survival**

PREDICTIVE VALIDITY COEFFICIENTS

To evaluate the utility of the LAB Test, and in particular the auding components of the LAB as a supplement to the AFQT for selection and classification, first-order and multiple correlations of LAB and AFQT scores with the criterion variables were computed. These coefficients of predictive validity are shown in Table 18. Since educational level is a factor in enlistment qualification decisions, years of school completed was included as an additional predictor in the computation of coefficients of predictive validity.

Table 18

Coefficients of Predictive Validity (r) for LAB and AFQT Scores

Predictor	Criterion: AFQT	Qualification Status[2]	Attrition[3] 1-6 Months	Attrition[3] 7-30 Months	Total Attrition[3]	Months Before Separation[4]	Highest Pay Grade Achieved
Maximum[1]	1.00	.74 (100)	–.59 (100)	–.62 (100)	–.69 (100)	1.00	1.00
LAB: Auding Paragraphs	.64	.51 (69)[5]	–.18 (31)	–.13 (21)	–.20 (29)	.13	.17
Reading Paragraph	.66	.53 (72)	–.17 (29)	–.08 (13)	–.17 (25)	.12	.13
A/R Vocabulary	.58	.53 (72)	–.15 (25)	<–.01 (<2)	–.10 (14)	.17	.07
Reading Vocabulary	.62	.56 (76)	–.16 (27)	<–.01 (<2)	–.11 (16)	.20	.11
Decoding Total	.62	.50 (68)	–.12 (20)	–.03 (5)	–.10 (14)	.12	.13
Total	.71	.59 (79)	–.18 (31)	–.07 (11)	–.16 (23)	.17	.15
AFQT Percentile	--	.63 (85)	–.11 (19)	–.11 (18)	–.15 (22)	<.01	.20
Education	.29	.19 (27)	–.09 (15)	–.07 (11)	–.11 (16)	<.01	.22
Best Combination	--	R. Vocab. + AFQT .67 (91)	LAB Tot + Educ. –.20 (34)	Aud Para + R. Vocab. –.15 (24)	Aud Para + Educ. –.23 (33)	R. Vocab + AFQT .23	AFQT + Educ. .27
Best Combination of AFQT + LAB	--	R. Vocab. + AFQT .67 (91)	(Aud Para + AFQT –.18(31)) LAB Tot + AFQT –.18 (31)	Aud Para + AFQT –.14 (23)	Aud Para + AFQT –.21 (30)	R. Voc. + AFQT .23	Aud Para + AFQT .22
Number Entering Into Correlations	2009	1993	664	560	627	135	704
Minimum Significant r .05	.06[6]	.06[6]	.09[7]	.09[7]	.09[7]	.17[8]	.09[7]
.01	.08[6]	.08[6]	.12[7]	.12[7]	.12[7]	.23[8]	.12[7]

[1] Maximum coefficient of predictive validity is less than 1.00 when the variable being predicted is a category (e.g., attrition) rather than continuous.

[2] Qualification status assignments were based on old AFQT norms.

[3] Only attrition for failure to meet minimum behavioral or performance criteria was predicted. See Table 15.

[4] For those who attrited.

[5] Numbers in parentheses express r's as percent of maximum r possible.

[6] Based on 1000 degrees of freedom.

[7] Based on 500 degrees of freedom.

[8] Based on 125 degrees of freedom.

Before discussing the coefficients presented in Table 18, it should be noted that the maximum possible correlation coefficients for predicting qualification status and attrition are considerably less than 1.00. This is due to the calculation of point-biserial correlations between the continuously distributed predictor variables and the dichotomized criterion variables (i.e., one is either qualified or not; attrited in six months or not, etc.). The percentage of the maximum correlation that was obtained for each predictor-criterion set is shown in parentheses next to each predictive validity coefficient.

Predicting Qualification Status

The military has traditionally made selection decisions based on the assumption that cognitive capacity, specifically the sort of abilities required by the ASVAB tests, is related to a successful service career. On this basis, minimum ASVAB scores are set by each service to screen out people who do not have the minimum requisite cognitive capacity for adequate performance. Although different services set different standards and educational level is also a factor affecting qualification for several of the services, the difference between ASVAB means for the qualified and unqualified groups in the present study (Table 17) indicates that the selection process is indeed separating out the lower performers on the ASVAB.

As expected, because it enters into the decision as to whether a person is or is not qualified for military service, AFQT is the most highly correlated cognitive test predictor with the qualification status criterion, with the correlation of .63 reaching 85 percent of the maximum possible value of .74.

The high correlations between AFQT and LAB scores, including Auding scores, given in Chapter 3, Table 13 are consistent with the position of the developmental model that much of the variation being tapped by all the tests given in this study involves a general languaging capacity, not specific to the reading modality alone. In light of these high correlations, it is not surprising that Table 18 shows that LAB scores have high coefficients of validity for predicting qualification status. However, it is striking that LAB scores should account for nearly as much of the variance in qualification as do the ASVAB scores on which selection decisions are based. That the correlation between auding score and qualification is so high indicates that service selection decision makers are concommittantly dividing the population according to auding ability when they set AFQT standards for qualification.

Predicting Attrition

While none of the cognitive tests or subtests are very efficient for predicting attrition within the first six months, the LAB Auding Paragraphs subtest is as efficient as the total LAB Battery with a correlation of −.18, which is 31 percent of the maximum correlation possible. In contrast, the AFQT correlation achieves a value that is only 19 percent of the maximum possible. For predicting attrition during the first six months of service, the Auding Paragraph subtest emerges as the LAB subtest best combining with AFQT to increase the predictive efficiency of the AFQT. In this case, the multiple correlation reaches −.18 which is not higher than the first order correlation of Auding Paragraph with a six-month attrition. Nonetheless, the predictive efficiency of the AFQT is increased three-fold.

Insofar as selection decisions have prevented from enlisting personnel who do not have the cognitive skills needed for job success, correlations between cognitive assessments and criteria indicating such success will be attenuated. In addition, the success of classification decisions, placing less able personnel in less demanding fields will also have the effect of lessening correlations when attrition is used as a criterion variable. These effects are also compounded by the fact that cognitive abilities are only one subtest of a number of factors, e.g., motivation or personality, affecting attrition. The latter factors may be especially important for attrition that is not due to communication skills, learning problems, or other cognitive abilities. The present analysis was unable to separate attrition due to cognitive problems from many other types of behavioral problems (see Table 16) due to limitations in the DMDC data base. In view of the foregoing, the correlations found between LAB scores and attrition measures in this study are substantial, particularly when it is remembered that the maximum correlations possible for these point-biserial criteria are considerably less than 1.00.

In earlier discussion it was hypothesized that since initial training requires more concentrated learning from oral and written language than does subsequent job performance, language skills should be more closely related to attrition in the initial period of service. This hypothesis is supported by the correlations of LAB and AFQT scores with attrition in the first six months of service in contrast with months seven through 30.

In light of the preceding discussion of the importance of auding performance to training and job success, it is encouraging to find that LAB Auding Paragraph scores are (1) the best predictor of all three types of attrition and (2) the LAB variable which leads to the best prediction when combined with AFQT. It was stated earlier that while reading and auding are equally important in formal training, in many jobs reading demands may be minimal, while listening skills are crucial. This idea is consistent with the relatively large difference between the predictive power of LAB Auding and Reading scores for attrition in months 7-30 after the majority of training is completed.

The other two columns of coefficients of Table 18 add supporting information regarding the usefulness of the LAB Auding Paragraphs score as a supplement to the AFQT. When months of service for attritees are considered, months before separation reflects early or late attrition, which is associated with the heavy literacy and learning demands of training versus job performance demands for these skills. Consistent with this situation is the fact that a test score known to correlate highly with school performance, vis., Reading Vocabulary, is the best predictor of months of service prior to attrition.

Predicting Paygrade

Taking notice of the fact that in this sample there is a restricted range for paygrade, reflecting minimum and maximum time in grade regulations, the data of Table 18 indicate that AFQT and education are the best predictors of paygrade. This doubtless reflects educational requirements for obtaining promotions. It is notable, however, that the LAB score adding most predictive power to AFQT is again Auding Paragraphs.

Chapter 5

SUMMARY, DISCUSSION AND CONCLUSIONS

The present study continues a program of research initiated by the Department of Defense (OASD/M&RA) in 1968 to better understand the nature of literacy, and the role of literacy in aptitude assessment, job training, and job performance in the armed services (Sticht, 1975). The value of research on literacy is apparent when it it realized that all of the subtests of the Armed Services Vocational Aptitude Battery (ASVAB) require the comprehension of written language; all military job technical skills training programs require the comprehension and production of written language presented in a wide variety of formats (textbooks; technical manuals; charts, tables, forms; end of course tests, etc.) and all military jobs are performed using a variety of written language tools that personnel must comprehend and produce (Sticht and Zapf, 1975; Curran, 1980). Additionally, success in both military and civilian environments requires that personnel be able to comprehend and produce written language and a wide variety of graphic tools for thinking and performing such as street signs, dispensary notices, finance forms, bus schedules, etc. (Hooke and Sticht, 1981).

A DEVELOPMENTAL MODEL OF LITERACY

Because literacy is such a ubiquitously necessary, though rarely sufficient information processing component in the performance of a wide range of military tasks, it is important to have as sound an understanding as possible of the nature of literacy. Chapter 2 of this report presents, in a highly summary form, a conceptual framework for understanding literacy that (1) defines various information processing skills involved in literacy; (2) discusses how such skills develop in relation to other skills, such as the oracy skills of auding and speaking; (3) describes how knowledges and skills develop over time through four stages that are related by the transformation and storage of information in memory and, (4) discusses how the latter information is used to develop new capabilities for communication and thinking. In the latter case, the model discusses the development of reading through the interaction of the skills and knowledges involved in visual information processing with the skills and knowledges involved in comprehending oral language, i.e., auding.

Drawing upon the work of Fries (1963), the developmental model of literacy distinguishes three components involved in the communication process: Knowledge (meaning; thoughts), language as an internal representation of knowledge, and speech or writing as external representations of the internal language signals. According to the developmental model, the ability to comprehend oral language, that is, to form internal language patterns from speech, and to use those patterns to construct knowledge (meaning; thoughts) develops first. Then, in the typical case, in learning to read, the person learns to use the same internal language system in response to written, graphic symbols, as is used in comprehending spoken language. In other words, auding and reading use the same, or very much the same internal language system for representing the same thoughts, that is, they share the same meaning system. Thus, in learning to read, people learn to comprehend by reading what they could previously comprehend by auding.

Based on this analysis, it is possible to consider that performance on an aptitude test, such as the Word Knowledge (WK) or Automotive Information (AI) subtests of the ASVAB may suffer if[1] (1) the relevant knowledge is not in the person's memory, (2) the knowledge is there but the internal (oral) language for representing the knowledge is missing; or (3) both the knowledge and internal language are available but the person cannot recode the external, graphic representation of language into the internal language, or cannot do this efficiently enough to satisfy task requirements, as in a speeded test. The third case constitutes the literacy (reading) component of the aptitude test task. Whenever a task includes the comprehension of written language, whether on a reading test, an aptitude test, an in-progress or end of training paper-and-pencil achievement test, or a job knowledge, paper-and-pencil proficiency test, literacy (reading) is a necessary, though not sufficient, component of task performance.

THE LITERACY ASSESSMENT BATTERY (LAB)

In the typical reading or aptitude test that requires reading, it is not possible to determine which of the three factors listed above, i.e., knowledge, oral language, or reading may operate to limit performance on the test. To partially overcome this limitation, the Literacy Assessment Battery (LAB) was developed to determine if a person who scores low on tests requiring reading actually has higher oral language skills and is, therefore, primarily unskilled in transforming the written language into the earlier developed internal oral language.

The LAB test consists of three subtests (Chapter 3). The Paragraphs Test consists of four paragraphs, two of which are presented by spoken language for auding, and two of which are presented by written language for reading. Questions are asked following each paragraph to indicate how well the materials can be recalled. By this means, comprehension and recall of oral language can be compared to comprehension and recall of written language.

The LAB Vocabulary Test consists of words taken from the Paragraphs Test to determine if word knowledge might be a factor limiting performance by either or both auding and reading in the Paragraphs Test. In the Auding/Reading Vocabulary Test the words from the Auding Paragraphs are presented both orally and in written language form. Here the intent is to give the person the choice of modalities because the main interest is to find out if word knowledge is known regardless of modality. The Reading Vocabulary Test is given in written language form. Thus, this test requires reading. By comparing the Auding/Reading and Reading Vocabulary Tests, one can gain an idea of the extent to which having the Auding option enhances Vocabulary Test performance.

The LAB Decoding Test is based directly on the concept from the developmental model that in reading the person transforms the graphic language into the same internal language form that results from auding. The more efficiently one can accomplish this transformation, the more competent the person is considered to be at decoding writing to internal oral language. To evaluate efficiency at decoding, the Decoding Test requires the person to aud a spoken message while at the same time reading a printed version of

[1] Additional factors may affect the performance of aptitude test and other literacy tasks, such as the amount of information that must be held in working memory to be used to draw inferences for making responses, etc., but these task demands are not considered here.

the message. Then, from time-to-time, a word being spoken differs from a word on the printed page, and the person must circle this mismatch. By accelerating the rate of speech, one can determine how efficiently the person can perform this simultaneous auding and reading task in which both oral and printed external forms of language must be converted into the same internal form of language so that a "same" or "different" judgement can be made. Poorer decoders will not be able to perform this task well at the faster rates of speech.

With the three subtests of the LAB, it is possible to determine whether a person can comprehend and remember connected discourse presented in spoken form better than in written form, whether the person knows the vocabulary of the paragraphs but cannot process connected discourse well, and whether or not the person is an efficient decoder.

NORMATIVE DATA FOR THE LAB

To determine the distribution of auding and reading skills in the population of military applicants, Mathews, Valentine, and Sellman (1978) administered the LAB test to over 4,500 applicants. The present research analyzed the LAB data obtained by Mathews et al. to develop normative data for the LAB, and to relate scores on the LAB to other literacy tests and composite scores of the ASVAB. Chapter 3 presents the details of this analysis. Major findings include:

(1) Comparison of auding and reading. Consistent with the developmental model it was found that:

- Auding and reading are highly correlated (.73 for LAB Auding and Reading Paragraphs) indicating that people who are unskilled at reading are also the least skilled in comprehending oral language.
- The lower the reading skill, the more likely one is to find people who comprehend better by auding than by reading; however, overall, across the full range of reading skills in the norming population, except at the very lowest levels, people tended to perform better by reading than by auding. This suggests that for the majority of poor readers who apply for military service, literacy (decoding) problems are accompanied by oracy (low internal oral language vocabulary and ability to process connected discourse) problems.
- Performance on the LAB Vocabulary Tests exceeded performance on LAB Paragraphs Tests suggesting that though knowledge of the vocabulary used to express the concepts presented in the Paragraphs Tests was fairly high in the norming population. Processing of connected discourse, as in lectures or textbooks, presents additional demands that make the task more difficult.
- Performance on the LAB Auding/Reading Vocabulary test exceeded performance on the LAB Reading Vocabulary test across all levels of reading skill. Analyses of a sample of answer sheets suggested that because the Auding/Reading method forced examinees to consider each item and them move along to the next, more people completed the Auding/Reading Vocabulary test than the Reading — only Vocabulary test. Thus, the method of information display, and not necessarily the modality of information presentation may have resulted in a more complete and accurate assessment of people's vocabulary knowledge with the result being higher Auding/Reading Vocabulary test scores across reading grade levels. This suggests that other vocabulary tests that rely solely on self-pacing, such as the Word Knowledge (WK) subtest of the ASVAB, may underestimate vocabulary knowledge and hence aptitude for learning of verbal materials.

- Performance on the LAB Decoding Test showed an average systematic decline as the rate of speech was increased, indicating that efficiency of decoding is, indeed, stressed by this technique. The correlation of the Decoding total score with Paragraph Reading was .77 while it was .63 with the Auding Paragraph test. This suggests a greater relationship to reading than to auding skills as would be expected if reading decoding is an additional process to auding.

(2) Correlations of the LAB With Other Literacy Tests. Tables 13 and 14 show intercorrelations among the LAB subtests and the subtests of the Gates-MacGinitie, Nelson-Denny, Basic Skills Assessment, and ASVAB tests. The data indicate that:

- Correlations of Space Perception, LAB Auding Paragraphs, and LAB Reading Paragraphs with other general literacy tests increase in that order, with average correlations of .35, .63, and .73, respectively. These data were interpreted as indicating that, consistent with the developmental model, skills develop from prelinguistic modes of information pick-up through looking and listening to form internal images; then, listening to speech leads to oral language; and finally, looking at written language combines visual information pick-up with oral languaging to produce reading. Thus, all reading assessment is simultaneously an assessment of space (imagery) perception and languaging as well as decoding of written symbols. According to this model, correlations among literacy tests ought to be highest when reading tests are related to other reading tests, next highest when oral language tests are correlated with reading tests; and least highest when a measure of the prelinguistic skills of looking and listening to form internal images is correlated with reading tests. This is the pattern of the correlations given in Table 14.
- Correlations of the Space Perception, LAB Auding Paragraphs, and LAB Reading Paragraphs tests with the special knowledge literacy tests of the ASVAB show increased correlations among Space Perception and the special knowledge tests, and decreased correlations of LAB Auding and Paragraph tests with the special knowledge tests in comparison to the general literacy tests. The differences in correlations among general and specific knowledge literacy tests suggest that people's relative rank order on a literacy test will depend, among other things, upon the particular body of knowledge being addressed by the test. Additional analyses indicate, for instance, that the Automotive Information (AI) subtest of the ASVAB correlates .73 with Shop Information (SI), which is what would be expected if related bodies of special knowledge are being addressed, via language and imagery mediators, in these two tests. This correlation holds even though AI correlates only .43 with the LAB Reading Paragraphs test, and at a similar value for the Gates-MacGinitie and Nelson-Denny reading tests. This suggests that it is the knowledge base, rather than the reading decoding skills or general language knowledge, such as use of function words (the, an, etc.) and syntax that is playing the major role in ordering examinees on the AI and SI tests. Here as the correlations of Table 14 suggest, space perception and oral and written language skills are absolutely necessary for test performance, yet they are clearly not sufficient. Well structured bodies of knowledge, sometimes called schemata (Rumelhart, 1980) appear to play a predominant role in the special knowledge literacy tests of the ASVAB.

PREDICTING SUCCESS IN THE MILITARY WITH THE LAB

Earlier research (Sticht, et al., 1970) suggested that assessment of oracy skills in addition to the literacy skills assessed in the AFQT might permit a more accurate identification of less literate applicants for military service who could, by virtue of their oral language skills, benefit from military training and perform satisfactorily on the job. Because the LAB assesses both oracy (auding) and literacy (reading) skills, it permits the evaluation of this hypothesis.

Chapter 4 presents the details of a study to evaluate the LAB as a potential supplement to the AFQT for predicting success in the military. Major findings are:

- Predicting qualification status. The LAB total score correlated almost as well with Qualification status (qualified vs not-qualified for military service) as did the AFQT which was actually used in determining the qualification status. The LAB total score accounted for 35% of the variance in the Qualification status criterion, while the AFQT accounted for 40% of this variance. Adding the LAB Reading Vocabulary subtest to the AFQT produced a multiple correlation that accounted for 45% of this variance. These correlations reflect, to a large extent, the fact that both the LAB and AFQT assess language and reading knowledge and skills (AFQT and LAB Total scores correlate .71).
- Predicting attrition. The prediction of attrition from military service was performed separately for attrition in the first six months, which reflects the learning demands of initial entry and job training and orientation, and months seven through 30, which reflects a period of job performance in which cognitive skills are stressed least. Additionally, months to attrition was used as a criterion for just that subpopulation which in fact attrited.
  - Attrition in the first six months. The LAB Auding Paragraphs test correlated –.18 with attrition in the first six months which was 31% of the maximum correlation of –.59 attainable with the point-biserial correlation analysis. AFQT correlated –.11 with six months attrition. The LAB subtest that emerged as the best addition to supplement the AFQT was LAB Auding Paragraphs, though the multiple correlation of –18 did not exceed the first order correlation of LAB Auding Paragraph with attrition.
  - Attrition in months 7-30. Again LAB Auding Paragraphs was the best LAB subtest predicting attrition in months 7-30, with an r of –.13 compared to an r of –.11 again for the AFQT. Also, the LAB Auding Paragraphs emerged as the best supplement to the AFQT and increased the r from –.11 to –.14 (maximum r equals –.62 using the point-biserial correlation analysis).
  - Predicting total attrition across 30 months. The LAB Auding Paragraphs once again was the best predictor of this criterion with a correlation of –20 compared to –.15 for the AFQT. Also, the LAB Auding Paragraphs test emerged as the best supplement to the AFQT and increased the correlation from –.15 to –.21, not much better than the LAB Auding Paragraphs test alone.
  - Predicting months to attrition. The Lab Reading Vocabulary subtest was the best cognitive test predictor of this continuous variable,

with a correlation of .20. The AFQT correlation with this variable was less than .01. The LAB Reading Vocabulary test emerged as the best supplement to the AFQT for predicting this criterion and increased the r from .01 to .23.

- Predicting promotion (paygrade achieved). Education level emerged as the best predictor of paygrade achieved, with an r of .22 compared to .20 for the AFQT. The LAB Auding Paragraphs test emerged as the best LAB subtest to supplement the AFQT and increased the r from .20 to .22.

(4) Correspondence between predictor and criterion variables. The foregoing supports the hypothesis stated earlier that adding oracy skills to the literacy skills assessed by the AFQT might improve the accuracy of selection and classification procedures. Whereas in no case are correlations very large, nonetheless, in four out of the five predictive validity evaluations, the LAB Auding Paragraphs subtest emerged as the best LAB subtest to supplement the AFQT.

The fact that attrition correlation coefficients were, in general, quite small, may reflect the lack of correspondence between the predictor and criterion variables. As Table 16 shows, attrition due to behavior problems includes a number of factors, such as fraudulent entry, that may have little base in cognitive task performance. The fact that, in the present study, the prediction of attrition in the first six months was more accurate than in months seven through 30 is consistent with the fact that the first six months of service makes heavier demands on oracy, literacy, and learning skills like those assessed in the LAB (Sticht, et al., 1976; Sachar and Duffy, 1977).

When general reading tests and aptitude tests that involve reading (e.g., the AFQT) were used to predict various job proficiency indicators in an earlier study (Sticht, 1972), it was found that correlations of reading and AFQT tests were highest for predicting performance on job reading task tests (r = .78–.64), next highest in predicting job knowledge, paper-and-pencil test scores (r = .57–.36); lower in predicting hands-on job performance test scores (r = .40–.30) and lowest for predicting supervisor ratings (r = .17–.06). The latter are within the range of values found in the present study for predicting attrition.

In general, the foregoing supports the "point-to-point" theory of Asher, which Vineberg and Joyner (1981, p. 31) explain considers that "Higher validities may be possible when common or similar elements are present in the predictor and criterion." Thus, when literacy is a component of both predictor and criterion variables, predictive validity should be higher than when literacy is not a common element in the two variables. Inasmuch as the present study used job performance criteria with an unknown, but presumably small point-to-point relationship to the oracy, literacy, and aptitude tests used, it is not surprising that, even though the LAB Auding Paragraphs test doubled or tripled the variance in attrition predictable by the AFQT, overall predictive validity coefficients were relatively small. A forthcoming report will examine the effects on the prediction of attrition through separate analyses of data for the different armed services.

CONCLUSIONS

Understanding the nature of literacy and the relationship of literacy to aptitude assessment can lead to improvements in selection and classification. The administration of the Literacy Assessment Battery (LAB) based on the developmental model of literacy described in Chapter 2 increased the validity of the AFQT two or threefold in predicting attrition from the military in the first 30 months. Continued exploration of the nature and interrelationships among oracy, literacy, and vocational aptitude tests is suggested by these results.

REFERENCES

REFERENCES

Brown, J., Nelson, M., and Denny, E. *The Nelson-Denny Reading Test. Examiner's Manual.* Boston, MA: Houghton Mifflin Company, 1976.

Canter, R. *Organization Management and the Volunteer Force: Policy and Research Issues.* Paper presented at the Security Issues Symposium: Army Strategic Environment 1985-2000. Strategic Studies Institute, U.S. Army War College, Carlisle Barracks, PA, April 1978.

Curran, T. (Ed.) · *Tri-Service Literacy and Readability: Workshop Proceedings.* NPRDC Special Report 80-12, San Diego, CA: Navy Personnel Research and Development Center, March 1980.

Foulke, E. and Sticht, T. Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin*, 1969, 72, 50-62.

Fox, W., Taylor, J. and Caylor, J. *Aptitude Level and the Acquisition of Skills and Knowledges in a Variety of Military Training Tasks.* HumRRO TR 69-6. Alexandria, VA: Human Resources Research Organization, May 1969.

Frederiksen, J. *A Componential Theory of Reading Skills and Their Interactions: Final Report.* Report No. 4649. Cambridge, MA: Bolt, Beranek, and Newman Inc., April 1981.

Fremer, J., Swineford, F., and Zieky, M. *Technical Manual: Basic Skills Assessment Program.* Menlo Park, CA: Addison-Wesley, 1978.

Fries, C. *Linguistics and Reading.* New York: Holt, Rinehart and Winston, 1963.

Gates, A. and MacGinitie, W. *Gates-MacGinitie Reading Tests, Survey D, Teacher's Manual.* New York: Teachers College Press, 1965.

Holden, C. Doubts mounting about all-volunteer force. *Science*, 1980, 209, 1095-1099.

Hooke, L. and Sticht, T. *Instructional Systems Design for the Army's On-Duty Education Program: Task 1: Development of Terminal Objectives for the Army's EXCEL Program.* Final Report Task 1. Alexandria, VA: Human Resources Research Organization, October 1981.

Marshalek, B. *Trait and Process Aspects of Vocabulary Knowledge and Verbal Ability.* TR No. 15. Stanford, CA: School of Education, Stanford University, May 1981.

Mathews, J., Valentine, L., and Sellman, W. *Prediction of Reading Grade Levels of Service Applicants from Armed Services Vocational Aptitude Battery (ASVAB).* AFHRL-TR-78-82. Brooks Air Force Base, TX: Air Force Human Resources Laboratory, December 1978.

Miller, E., Nystrom, C., and Hicks, J. *Reading Ability and Other Correlates of the SQT Written Component.* HumRRO FR-MTRD(TX)-80-1. Alexandria, VA: Human Resources Research Organizatión, March 1980.

Mitzel, M. The functional reading word list for adults. *Adult Education.* 1966, 67-69.

OASD(MRA&L). *Implementation of New Armed Services Vocational Aptitude Battery and Actions to Improve the Enlistment Standards Process.* Washington, D.C.: A Report to the House and Senate Committees on Armed Services, December 31, 1980.

OASD(MRA&L). *Aptitude Testing of Recruits.* Washington, D.C. A Report to the House Committee on Armed Services, July 1980.

Pelligrino, J. and Glaser, R. Components of inductive reasoning. In: R. Snow, P.A. Federico, and W. Montague (Eds.), *Aptitude, Learning, and Instruction: Cognitive Process Analysis of Aptitude. (Vol. 2).* Hillsdale, N.J.: Lawrence Erlbaum, 1980.

Rumelhart, D. Schemata: The building blocks of cognition. In: R. Spiro, B. Bruce, and W. Brewer (Eds.) *Theoretical Issues in Reading Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum, 1980.

Sacher, J. and Duffy, T. *Reading Skill and Military Effectiveness.* Paper presented at the Ninth Annual Convention of the American Educational Research Association, 1977.

Sellman, W. and Valentine L. *Aptitude Testing, Enlistment Standards, and Recruit Quality.* Paper presented at the meetings of the American Psychological Association, August 1981.

Sternberg, R. *Intelligence, Information Processing, and Analogical Reasoning: The Componential Analysis of Human Abilities.* Hillsdale, N.J.: Lawrence Erlbaum, 1977.

Sticht, T. Learning by Listening. In: R. Freedle and J. Carroll (Eds.) *Language Comprehension and the Acquisition of Knowledge.* Washington, D.C.: V.H. Winston & Sons, 1972.

Sticht T. (Ed.) *Reading for Working.* Alexandria, VA: Human Resources Research Organization, 1975.

Sticht, T. Comprehending reading at work. In: M. Just and P. Carpenter (Eds.) *Cognitive Processes in Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum, 1977.

Sticht, T. The Acquisition of Literacy by Children and Adults. In: F. Murray and J. Pikulski (Eds.) *The Acquisition of Reading.* Baltimore, MD: University Park Press, 1978.

Sticht, T. and Zapf, D. *Reading and Readability Research in the Armed Services.* FR-WD-CA-76-4. Alexandria, VA: Human Resources Research Organization, September 1976.

Sticht, T. and Beck, L. *Experimental Literacy Assessment Battery (LAB).* AFHRL-TR-76-51. Brooks Air Force Base, TX: Air Force Human Resources Laboratory, August 1976.

Sticht, T. Caylor, J., Kern, R., and Fox, L. *Determination of Literacy Skill Requirements in Four Military Occupational Specialties.* HumRRO TR 71-23. Alexandria, VA: Human Resources Research Organization, November 1971.

Sticht, T., Fox, L., Hauke, R., and Zapf, D. *Reading in the Navy.* HumRRO FR-WD-CA-76-14. Alexandria, VA: Human Resources Research Organization, September 1976.

Sticht, T., Beck, L., Hauke, R., Kleiman, G., and James, J. *Auding and Reading: A Developmental Model.* Alexandria, VA: Human Resources Research Organization, 1974.

Thorndike, R. (Ed.) *Educational Measurement.* Washington, D.C.: American Council on Education, 1971.

Vineberg, R. and Joyner, J. *Prediction of Job Performance: Review of Military Studies.* FR-MTRD-(CA)-81-6. Alexandria, VA: Human Resources Research Organization, March 1981.

Vineberg, R., Sticht, T., Taylor, E., and Caylor, J. *Effects of Aptitude (AFQT), Job Experience, and Literacy on Job Performance: Summary of HumRRO Work Units UTILITY and REALISTIC.* TR 71-1. Alexandria, VA: Human Resources Research Organization, February 1971.

Waters, B. *The Test Score Decline: A Review and Annotated Bibliography.* Technical Memorandum 81-2. Washington, D.C.: Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics), Directorate for Accession Policy, August 1981.

Appendix A

NORMING DATA FOR THE LITERACY ASSESSMENT BATTERY (LAB)

This appendix contains tables for converting LAB raw scores into percentile scores, reading grade level scores on three reading tests, and ASVAB composites (GT and AFQT). See Chapter 3 of this report for instruction in using these tables. Further information is given in: T. Sticht, L. Hooke, and J. Caylor. *Manual for the Administration and Interpretation of the Literacy Assessment Battery (LAB).* Final Report 81-10. HumRRO FR-ETSD-81-10. Alexandria, VA: Human Resources Research Organization, May 1981.

Table A-1

Norms for Converting Auding Paragraphs Raw Scores Into Reading Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

Auding Paragraphs Raw Score	Percentile	Gates-MacGinitie RGL	Nelson Denny RGL	BSA	AFQT Percentile	GT Percentile
0	1	2.2			1	1
1	2	2.5	6.1		3	2
2	4	2.9	6.3		4	5
3	5	3.2	6.4		5	6
4	6	3.5	6.5		6	7
5	8	3.9	6.6		7	8
6	10	4.4	6.7		8	9
7	12	4.7	6.9	100	10	10
8	15	5.0	7.0	105	11	11
9	19	5.4	7.2	112	12	13
10	23	5.8	7.3	119	13	14
11	27	6.3	7.6	126	15	16
12	33	6.9	7.8	133	18	19
13	38	7.5	8.2	142	22	22
14	45	8.2	8.7	149	27	29
15	51	9.1	9.2	153	32	35
16	59	9.9	10.0	160	41	44
17	69	10.7	11.1	165	52	55
18	76	11.1	12.2	170	50	64
19	82	11.6	13.1	172	67	70
20	89	11.7	13.9	176	78	78
21	94	11.9	14.7	179	86	84
22	97		14.9	179	91	90
23	99		15.0	181	97	95
24	99				99	99

Table A-2

Norms for Converting Reading Paragraphs Raw Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

Reading Paragraphs Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
0	1	2.2			1	1
1	3	2.7	6.1		3	5
2	4	2.9	6.3		4	6
3	5	3.3	6.5		5	7
4	7	3.7	6.6		6	9
5	9	4.1	6.7		8	9
6	10	4.4	6.8		9	9
7	12	4.7	6.9	100	10	10
8	14	4.9	7.0	103	11	11
9	18	5.3	7.1	110	12	12
10	21	5.7	7.3	114	14	13
11	24	6.0	7.4	121	15	15
12	29	6.4	7.6	128	17	17
13	33	6.9	7.9	135	18	19
14	37	7.4	8.1	140	21	22
15	42	7.9	8.5	144	25	26
16	48	8.5	8.9	151	30	31
17	54	9.4	9.5	156	35	38
18	60	10.0	10.0	160	41	45
19	69	10.7	11.2	165	53	55
20	77	11.2	12.3	170	61	65
21	85	11.7	13.5	174	72	73
22	93	11.8	14.4	179	84	84
23	97	11.9	14.9	179	91	90
24	99		15.0	181	99	99

Table A-3

Norms for Converting Auding/Reading Vocabulary Raw Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

Auding/Reading Vocabulary Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
0						
1						
2						
3						
4						
5						
6						
7		2.1				
8		2.2				
9		2.2				
10	1	2.3			1	1
11	2	2.5			2	3
12	3	2.7	6.1		3	5
13	4	3.0	6.4		4	5
14	5	3.3	6.5		5	6
15	7	3.7	6.6		6	7
16	8	3.9	6.6		7	8
17	9	4.2	6.7		8	9
18	11	4.5	6.8		10	9
19	13	4.8	6.9	100	11	10
20	15	4.9	7.0	105	11	11
21	17	5.2	7.1	107	12	12
22	20	5.5	7.2	114	13	13
23	24	5.9	7.4	119	15	15
24	30	6.6	7.7	130	17	17
25	40	7.7	8.3	142	23	24
26	52	9.2	9.3	156	33	36
27	76	11.2	12.1	170	60	64
28	99	11.9	15.0	181	99	99

Table A-4

Norms for Converting LAB Reading Vocabulary Raw Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

Reading Vocabulary Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
0						
1						
2						
3		2.1				
4		2.2				
5	1	2.3			1	1
6	2	2.5			2	3
7	3	2.7	6.1		3	5
8	4	2.9	6.3		4	5
9	4	3.2	6.4		4	5
10	5	3.5	6.5		5	6
11	6	3.6	6.6		5	6
12	8	3.9	6.6		7	8
13	9	4.1	6.7		8	9
14	10	4.4	6.7		9	9
15	11	4.5	6.8		10	9
16	12	4.7	6.9	100	10	10
17	13	4.8	6.9	103	11	10
18	16	5.0	7.0	105	12	11
19	17	5.2	7.1	107	12	12
20	20	5.5	7.2	114	13	14
21	23	5.8	7.3	119	15	15
22	27	6.2	7.5	126	16	16
23	32	6.8	7.8	133	18	18
24	39	7.6	8.2	142	23	23
25	48	8.7	9.0	151	31	33
26	62	10.2	10.2	163	44	48
27	82	11.6	13.1	172	67	70
28	99	11.9	15.0	181	99	99

Table A-5

Norms for Converting LAB Decoding Raw Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

Decoding Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
Decoding 100 WPM						
0	2	2.6	6.1		2	3
1	3	2.9	6.3		3	5
2	4	3.2	6.4		4	5
3	5	3.3	6.5		5	6
4	6	3.6	6.5		5	7
5	8	3.9	6.6		7	8
6	10	4.5	6.7	100	8	9
7	15	4.9	7.0	105	11	11
8	23	5.8	7.3	119	15	15
9	41	7.8	8.4	144	24	25
10	99	11.9	15.0	181	99	99
Decoding 150 WPM						
0	3	2.7	6.1		3	5
1	5	3.3	6.5		5	6
2	6	3.7	6.6		5	6
3	9	4.1	6.7		7	9
4	11	4.5	6.8	100	8	9
5	14	4.8	6.9	103	11	10
6	18	5.7	7.1	110	12	12
7	24	5.9	7.1	119	15	15
8	34	7.1	8.0	135	19	20
9	56	9.6	9.6	158	37	42
10	99	11.9	15.0	181	99	99
Decoding 200 WPM						
0	5	3.3	6.5		5	6
1	8	4.1	6.7		7	8
2	13	4.8	6.9	100	9	10
3	17	5.2	7.1	107	11	11
4	23	5.8	7.3	119	15	14
5	29	6.4	7.7	128	17	17
6	37	7.4	8.1	140	21	22
7	47	8.5	8.9	151	29	30
8	63	10.2	10.1	163	45	49
9	82	11.6	13.1	172	68	70
10	99	11.9	15.0	181	99	99
Decoding 250 WPM						
0	13	4.7	6.9	100	10	10
1	22	5.8	7.3	117	14	14
2	31	6.7	7.3	130	18	18
3	[illegible]	7.7	8.3	142	24	24
4	[illegible]	8.7	9.0	151	30	32
5	[illegible]	9.3	9.9	158	38	44
6	[illegible]	10.6	10.9	165	52	54
7	[illegible]	11.2	12.1	170	60	64
8	84	11.6	12.2	174	71	73
9	93	11.8	14.5	179	84	84
10	99	11.9	15.0	181	99	99

Table A-6

Norms for Converting LAB Decoding Total Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentages

Decoding Total Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
0	1	2.3			1	1
1-2	2	2.6	6.1		3	3
3-4	4	3.0	6.2		4	6
5-6	5	3.2	6.4		5	6
7-8	6	3.6	6.5		5	7
9-10	7	3.8	6.6		6	7
11-12	9	4.1	6.6		8	9
13-14	11	4.5	6.7	100	9	10
15-16	13	4.7	6.9	103	10	10
17-18	15	5.0	7.0	105	11	11
19-20	19	5.4	7.2	112	13	13
21-22	22	5.8	7.3	117	14	14
23-24	27	6.2	7.5	126	16	16
25-26	33	6.9	7.8	135	18	19
27-28	40	7.6	8.3	142	23	24
29-30	47	8.4	8.8	151	29	30
31-32	56	9.6	9.7	158	37	42
33-34	67	10.6	10.9	165	50	53
35-36	79	11.4	12.6	172	63	67
37-38	90	11.8	14.1	176	80	79
39-40	99	11.9	15.0	181	99	99

Table A-7

Norms for Converting LAB Total Raw Scores Into Reading Test Grade Level Equivalents, ASVAB Composite Equivalents and Percentiles

LAB Total Raw Score	Percentile	Gates-MacGinitie RGL	Nelson-Denny RGL	BSA	AFQT Percentile	GT Percentile
0-20	1	2.1				
21-25	1	2.2			1	1
26-30	2	2.4			2	2
31-35	2	2.6	6.1		3	2
36-40	3	2.8	6.3		3	5
41-45	4	3.1	6.4		4	5
46-50	6	3.5	6.5		5	6
51-55	7	3.8	6.6		6	7
56-60	8	4.1	6.6		7	8
61-65	10	4.3	6.7		9	9
66-70	11	4.6	6.8		10	10
71-75	13	4.8	6.9	100	11	10
76-80	15	5.0	7.0	105	11	11
81-85	18	5.2	7.1	110	13	12
86-90	22	5.8	7.2	117	14	14
91-95	25	5.9	7.4	121	15	16
96-100	29	6.5	7.7	128	17	17
101-105	35	7.1	8.0	137	19	21
106-110	42	7.9	8.5	144	25	26
111-115	49	8.6	8.9	151	30	32
116-120	57	9.7	9.8	158	33	43
121-125	68	10.6	10.9	165	52	54
126-130	79	11.4	12.6	172	62	67
131-135	90	11.7	14.1	176	80	79
136-140	97	11.9	15.0	179	91	90
141-144	99			181	99	99

Table A-8

Norms for Converting Gates-MacGinitie Reading Levels to Levels on ABLE and TABE Tests

Gates-MacGinitie	ABLE	TABE
4.0	5.5	
4.1	5.6	
4.5	5.7	
4.6	5.8	
4.8	6.0	
5.1	6.2	
5.3	6.3	
5.4	6.5	
5.6	6.6	
5.7	6.7	
5.8	6.8	8.0
6.0	6.9	8.1
6.1	7.0	8.1
6.2	7.2	8.1
6.3	7.3	8.2
6.4	7.4	8.3
6.6	7.5	8.4
6.7	7.6	8.5
7.0	7.7	8.9
7.2	7.8	9.0
7.4	7.9	9.1
7.5	8.0	9.2
7.7	8.1	9.4
8.1	8.2	9.5
8.4	8.3	9.6
8.7	8.5	9.7
8.9	8.6	9.8
8.9	8.6	9.8
9.0	8.7	9.8
9.3	8.8	9.9
9.5	8.9	9.9
9.7	9.0	10.0
10.0	9.2	10.1
10.4	9.4	10.2
10.6	9.5	10.3
10.8	9.6	10.4
11.0	9.8	10.5
11.3	10.1	10.8
11.5	10.3	11.2
11.7	10.4	11.3
11.8+	10.5+	11.4+